

(FILE 'USPAT' ENTERED AT 10:29:32 ON 30 MAY 96)

L1 0 S MEGAKARYOCYTE(W) KINASE
L2 0 S MEGAKARYOCYTE(5A) KINASE
L3 3 S MEGAKARYOCYTE(P) KINASE

=> d 1-3

1. 5,491,242, Feb. 13, 1996, Protein kinase C inhibitors; James R. Gillig, et al., 548/455 [IMAGE AVAILABLE]

2. 5,481,003, Jan. 2, 1996, Protein kinase C inhibitors; James R. Gillig, et al., 548/455, 312.1 [IMAGE AVAILABLE]

3. 5,378,464, Jan. 3, 1995, Modulation of inflammatory responses by administration of GMP-140 or antibody to GMP-140; Rodger P. McEver, 424/143.1; 514/8 [IMAGE AVAILABLE]

L4 0 Non-receptor (5A) Tyrosine^(w) Kinase

Set	Items	Description
S1	0	MEGAKARYOCYTE (W) KINASE
S2	5	MEGAKARYOCYTE (3N) KINASE
S3	67	(NON (W) RECEPTOR) (5N) (TYROSINE (W) KINASE)
S4	1	S3 AND MEGAKARYOCYTE

?d s2/3/all

Display 2/3/1

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1996 Knight-Ridder Info. All rts. reserv.

09606415 96128015

Transcription, splicing and editing of plastid RNAs in the nonphotosynthetic plant *Epifagus virginiana*.

Ems SC; Morden CW; Dixon CK; Wolfe KH; dePamphilis CW; Palmer JD

Department of Biology, Indiana University, Bloomington 47405, USA.

Plant Mol Biol (NETHERLANDS) Nov 1995, 29 (4) p721-33, ISSN 0167-4412

Journal Code: A60

Contract/Grant No.: GM-35087, GM, NIGMS

Languages: ENGLISH

Document type: JOURNAL ARTICLE

- end of record -

?

Display 2/3/2

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1996 Knight-Ridder Info. All rts. reserv.

09308492 95238492

The MATK tyrosine kinase interacts in a specific and SH2-dependent manner with c-Kit.

Jhun BH; Rivnay B; Price D; Avraham H

Department of Medicine, Deaconess Hospital, Harvard Medical School, Boston, Massachusetts 02215, USA.

J Biol Chem (UNITED STATES) Apr 21 1995, 270 (16) p9661-6, ISSN

0021-9258 Journal Code: HIV

Contract/Grant No.: HL51456, HL, NHLBI; HL46668, HL, NHLBI

Languages: ENGLISH

Document type: JOURNAL ARTICLE

- end of record -

?

Display 2/3/3

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1996 Knight-Ridder Info. All rts. reserv.

09200565 95130565

Structural and functional studies of the intracellular tyrosine kinase
MATK gene and its translated product.

Avraham S; Jiang S; Ota S; Fu Y; Deng B; Dowler LL; White RA; Avraham H
Division of Hematology/Oncology, New England Deaconess Hospital, Harvard
Medical School, Boston, Massachusetts 02215.

J Biol Chem (UNITED STATES) Jan 27 1995, 270 (4) p1833-42, ISSN
0021-9258 Journal Code: HIV

Contract/Grant No.: R01 HL51456, HL, NHLBI; R01 HL46668, HL, NHLBI

Languages: ENGLISH

Document type: JOURNAL ARTICLE

- end of record -

?

Display 2/3/4

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1996 Knight-Ridder Info. All rts. reserv.

09176341 95106341

Identification and characterization of Batk, a predominantly
brain-specific non-receptor protein tyrosine kinase related to Csk.

Kuo SS; Moran P; Gripp J; Armanini M; Phillips HS; Goddard A; Caras IW
Department of Neurobiology, Genentech, Inc., South San Francisco,
California 94080.

J Neurosci Res (UNITED STATES) Aug 15 1994, 38 (6) p705-15, ISSN
0360-4012 Journal Code: KAC

Languages: ENGLISH

Document type: JOURNAL ARTICLE

- end of record -

?

Display 2/3/5

DIALOG(R)File 155:MEDLINE(R)

(c) format only 1996 Knight-Ridder Info. All rts. reserv.

08802408 94117408

Identification and characterization of a novel tyrosine kinase from
megakaryocytes.

Bennett BD; Cowley S; Jiang S; London R; Deng B; Grabarek J; Groopman JE;
Goeddel DV; Avraham H

Division of Hematology/Oncology, New England Deaconess Hospital, Harvard
Medical School, Boston, Massachusetts 02215.

J Biol Chem (UNITED STATES) Jan 14 1994, 269 (2) p1068-74, ISSN
0021-9258 Journal Code: HIV

Contract/Grant No.: HL33774, HL, NHLBI; HL42112, HL, NHLBI; HL43510, HL,
NHLBI; +

Languages: ENGLISH

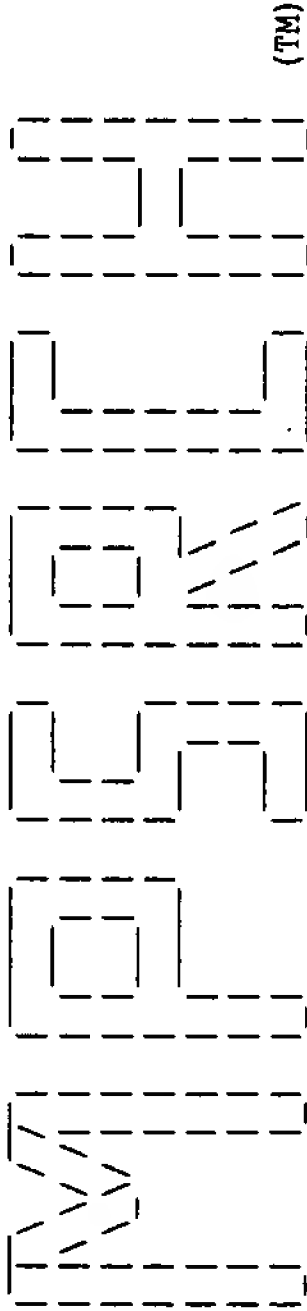
Document type: JOURNAL ARTICLE

- end of display -

May 31 09:25

US-08-426-509-6.rpr

1



(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri May 31 09:32:36 1996; MasPar time 21.94 Seconds

588.464 Million cell updates/sec

Tabular output not generated.

Title: >US-08-426-509-6

Description: (1-511) from US08426509.pep

Perfect Score: 3821

Sequence: 1 MYSWCMSNICQRLWEYLEP.....KLEDYFETDSSYSDANNFIR 511

Scoring table: PAM 150

Gap 11

Searched: 82306 seqs, 25270970 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: pir46

1:ann1 2:ann2 3:ann3 4:unann1 5:unann2 6:unann3 7:unann4

8:unann5 9:unann6 10:unann7 11:unann8 12:unrev1

13:unrev2

Statistics: Mean 48.603; Variance 117.818; scale 0.413

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	Length	DB	ID	Description	Pred. No.
1	2032	53.2	505	4	S24550		protein-tyrosine kin	0.00e+00
2	2026	53.0	506	4	S24553		protein-tyrosine kin	0.00e+00
3	1801	47.1	509	1	TVHAST		protein-tyrosine kin	6.43e-281
4	1754	45.9	536	10	S33569		protein-tyrosine kin	1.02e-272
5	1754	45.9	536	4	S29626		protein kinase yrk (1.02e-272
6	1745	45.7	362	4	S24551		protein-tyrosine kin	3.80e-271
7	1734	45.4	537	1	TVHUSY		protein-tyrosine kin	3.15e-269
8	1730	45.3	537	4	A43806		protein-tyrosine kin	1.57e-268
9	1726	45.2	537	1	TVHUSR		protein-tyrosine kin	7.82e-268
10	1685	44.1	512	4	A39719		protein-tyrosine kin	1.10e-260
11	1680	44.0	512	1	TVHULY		protein-tyrosine kin	8.15e-260
12	1682	44.0	532	4	B34104		protein-tyrosine kin	3.65e-260

May 31 09:25

US-08-426-509-6.rpr

2

13	1673	43.8	532	4	A34104	protein-tyrosine kin	1.35e-258
14	1674	43.8	533	1	TVCHS	protein-tyrosine kin	9.05e-259
15	1664	43.5	507	4	A39939	protein-tyrosine kin	5.00e-257
16	1662	43.5	587	1	TVFVPR	protein-tyrosine kin	1.11e-256
17	1660	43.4	505	1	TVHUHC	protein-tyrosine kin	2.49e-256
18	1659	43.4	517	4	A43807	protein-tyrosine kin	3.71e-256
19	1659	43.4	517	11	S24547	protein-tyrosine kin	3.71e-256
20	1657	43.4	568	1	TVFVSI	protein-tyrosine kin	8.28e-256
21	1652	43.2	557	1	TVFVS2	protein-tyrosine kin	6.15e-255
22	1647	43.1	509	4	A23639	protein-tyrosine kin	4.57e-254
23	1645	43.1	529	1	TVHUFR	protein-tyrosine kin	1.02e-253
24	1646	43.1	542	1	TVHUSC	protein-tyrosine kin	6.82e-254
25	1644	43.0	541	4	A43610	protein-tyrosine kin	1.52e-253
26	1639	42.9	503	1	TVMSHC	protein-tyrosine kin	1.13e-252
27	1639	42.9	503	4	A39973	protein-tyrosine kin	1.13e-252
28	1640	42.9	537	4	A45501	protein-tyrosine kin	7.56e-253
29	1638	42.9	541	11	S31645	protein-tyrosine kin	1.69e-252
30	1635	42.8	503	4	JQ1321	protein-tyrosine kin	5.61e-252
31	1636	42.8	541	1	TVCHYS	protein-tyrosine kin	3.76e-252
32	1635	42.8	543	1	TVHUYS	protein-tyrosine kin	5.61e-252
33	1627	42.6	509	1	OKHULK	protein-tyrosine kin	1.39e-250
34	1625	42.5	503	11	S18974	protein-tyrosine kin	3.09e-250
35	1622	42.4	526	1	TVFV60	protein-tyrosine kin	1.03e-249
36	1617	42.3	534	4	A44991	protein-tyrosine kin	7.63e-249
37	1610	42.1	528	1	TVFVVG9	protein-tyrosine kin	1.26e-247
38	1605	42.0	499	4	A40092	protein-tyrosine kin	9.35e-247
39	1598	41.8	505	11	S51647	protein-tyrosine kin	1.54e-245
40	1596	41.8	534	4	S33568	protein-tyrosine kin	3.44e-245
41	1591	41.6	526	4	S15582	protein-tyrosine kin	2.55e-244
42	1589	41.6	526	1	TVFVR	protein-tyrosine kin	5.68e-244
43	1589	41.6	545	7	S52313	pp62v protein - Rous	5.68e-244
44	1588	41.6	546	7	S52314	pp62v protein - Rous	8.48e-244
45	1584	41.5	526	13	S32774	protein-tyrosine kin	4.21e-243

ALIGNMENTS

RESULT	1	S24550	#type complete
ENTRY		protein-tyrosine kinase (EC 2.7.1.112)	1 - freshwater sponge
TITLE		(Spongilla lacustris)	
ALTERNATE_NAMES		src-type tyrosine kinase 1	
ORGANISM		#formal_name Spongilla lacustris	
DATE		07-May-1993 #sequence_revision 07-May-1993	#text_change
ACCESSIONS		28-Oct-1994	
REFERENCE		S24550	
#authors		Raulf, F.	
#submission		submitted to the EMBL Data Library, September 1991	
#accession		S24550	
#molecule_type		mRNA	
#residues		1-505	#label RAU
#cross-references		EMBL:X61601	
GENETICS			
#gene		srk1	
CLASSIFICATION		#superfamily protein-tyrosine kinase src; protein kinase	
KEYWORDS		homology; SH2 homology; SH3 homology	
FEATURE		ATP; phosphotransferase; tyrosine-specific protein kinase	
61-111		#domain SH3 homology #label SH3\	
122-214		#domain SH2 homology #label SH2\	
238-496		#domain protein kinase homology #label KIN\	
246-254		#region nucleotide-binding motif\	
268		#active_site Lys #status predicted	


```
FEATURE
66-115      #domain SH3 homology #label SH3\
126-218      #domain SH2 homology #label SH2\
238-497      #domain protein kinase homology #label KIN\
246-254      #region protein kinase ATP-binding motif\
268          #active_site lys #status predicted\
390          #binding_site phosphate (Tyr) (covalent) (by
              autophosphorylation) #status predicted
SUMMARY      #length 509 #molecular-weight 56885 #checksum 8721

Query Match      47.1%; Score 1801; DB 1; Length 509;
Best Local Similarity 53.8%; Pred.No. 6.43e-281;
Matches 245; Conservative 92; Mismatches 102; Indels 16; Gaps 10;

Db 64 fvalydyearisedlsfkkger-lqiintadgdwmyarslit-n--se---gyipstyv 115
|||:||||:| :||||: |:|:|:| ||:| | : | : |||| |
Qy 52 FVALFDYQARTAEDLSFRAGDKKLQVLDTLHEGWWFARHLEKRRDGSQQIQGYIPSNYV 111

Db 116 apeksyeaeewyfgdvkraeaeakrlmvriglpsgtflirkaetavgnfslsvrdgdsvkhy 175
| :| :|| | :| :|:|:| :|:|:|:| :|:|:|:| :|:|:|:| | |||| |
Qy 112 AEDRSLOAEPWFFGAIGRSDAEKQLLYSENKTSGLFIRESESQGEFSLSVLDGAVVKHY 171

Db 176 rvrkldtgggyfittrapfnslsylvqhytkdadglvcaltlpcpk-dkpvttggia-k--d 231
|:|:| | | :| :| :| :| | | | | | | | | | | | | | | :| :| :| |
Qy 172 RIKRLDEGGFFLTRRRIFSTLNEFVSHYTKTSDGLCVKLGKPCIKIQVPAPFDLSYKTV 231

Db 232 aweipreslrlnrklgagfgewagvwnnttqvavktlkgptmtpasfldaagvmkklr 291
||| | :| :| :| :|:|:|:| :|:|:|:| | |||| | | | | | :|:| |
Qy 232 QWEIDRNSIQLLRKRLSGQGFGVWEGLMNNTTPVAVKTLKPGSMDPNDFLREAQIMKNLR 291

Db 292 hkhlvqlyaicdsrepivivteymsggsildylskgegvnlqlptlidmaaqvasgmabl 351
| | :|:|:|:| :|:|:|:| | | | :| :| :| :| :| :| :| :| | | | | | |
Qy 292 HPKLIQYAVCTLEDPIYII TELMRHGSLOEYLQNDTGSKIHLTQQYDMAAQVASGMAYL 351

Db 352 eaqgyihrdlaarnilvgenyickvadfglarl--iedde-yaahgakfpikwtapeaa 408
| :| :|:|:|:|:|:| | | | | | | | :|:|:| :| :| :| :| :| :| :| | |
Qy 352 ESRNYIHRDLAARNVLVGEHNIYKVADFGLARVKVDNEDIYESRHEIKLPVKWTAPEAI 411

Db 409 lynrfctiksdvwsfgilmaeivtkgripypgmtnaqtiaevkegyrmpimpqcpeplyni 468
| :|:|:|:|:|:|:|:|:| :|:| :|:|:|:| | | :| :| :| :| :| :| :| :|
Qy 412 RSNKFSIKSDVWSFGILLYEIIITYGKMPYSGMTGAQVIQMLAQNRYRLPQPSNCPQFYNI 471

Db 469 mlqtwkdkpenrptfdylqgvledyfvsteggyrd 503
|| :| :| :| :|:|:| :| :| | | | | | | | :| |
Qy 472 MLECWNAEPKERPTFETLRWKLEDYF-ETDSSYS 505

RESULT 4
ENTRY      S33569 #type complete
TITLE      protein-tyrosine kinase (EC 2.7.1.112) yrk - chicken
ORGANISM   #formal_name Gallus gallus #common_name chicken
DATE       08-Dec-1993 #sequence_revision 03-Aug-1995 #text_change
ACCESSIONS S33569
REFERENCE   S33568
#authors    Sudol, M.; Greulich, H.; Newman, L.; Sarkar, A.; Sukegawa,
              J.; Yamamoto, T.
#journal     Oncogene (1993) 8:823-831
#title       A novel Yes-related kinase, Yrk, is expressed at elevated
              levels in neural and hematopoietic tissues.
#accession  S33569
##status     preliminary
##molecule_type mRNA
##residues   1-536 ##label SUD

CLASSIFICATION #superfamily protein-tyrosine kinase src; protein kinase
                homology; SH2 homology; SH3 homology
```

```
##cross-references EMBL:X67786
CLASSIFICATION #superfamily SH3 homology
KEYWORDS       phosphotransferase
FEATURE
88-137      #domain SH3 homology #label SH3
SUMMARY      #length 536 #molecular-weight 60002 #checksum 1930

Query Match      45.9%; Score 1754; DB 10; Length 536;
Best Local Similarity 53.1%; Pred.No. 1.02e-272;
Matches 246; Conservative 90; Mismatches 108; Indels 19; Gaps 9;

Db 86 fialydyearteddlsfqkgek-fhiinntegdwearsls---sgat----gyipsnyv 137
|:|:|:|:|:| | | | :| :|:|:| | | | | | :| :| :| | | | | |
Qy 52 FVALFDYQARTAEDLSFRAGDKKLQVLDTLHEGWWFARHLEKRRDGSQQIQGYIPSNYV 111

Db 138 apvdsiaqeeewyfgkigrkdaergllchncrgtfliresettkgavslsirdwdeakgd 197
| | :|:|:| | :| | | | | | | | | :|:|:|:|:| | :|:|:| | |
Qy 112 AEDRSLOAEPWFFGAIGRSDAEKQLLYSENKTSGLFIRESESQGEFSLSVLD-----GA 166

Db 198 hvkhykirklldsggyyitttraqfdtiqlvqhyieraaglcrlavpcpk-gtpkladls 256
|||:|:|:| | |:|:| | | :| :| :| :| :| :| :| :| :| | | |
Qy 167 VVKHYRIKRLDEGGFFLTRRRIFSTLNEFVSHYTKTSDGLCVKLGKPCIKIQVPAPFDLS 226

Db 257 vktkdvwepreslqlqlkgnqgfgevwmgtwngttkvavktlkgptmspeafleeaqi 316
|| | | | | :|:|:|:| :|:|:|:|:| | | | | | | | | | | | | :| | | | |
Qy 227 YKTVDQWEIDRNSIQLLRKRLSGQFGVWEGLMNNTTPVAVKTLKPGSMDPNDFLREAQI 286

Db 317 mkrlrhdklvqlyavvs--eepiyivtefmsqgsildflkdgdgrylklpqlvdmacaqiaa 375
|| | | | | | | | :|:|:|:|:| | | | | | | | :| :| :| :| :| :| :| :|
Qy 287 MKNLRHPKLIQYAVCTLEDPIYII TELMRHGSLOEYLQNDTGSKIHLTQQYDMAAQVAS 346

Db 376 gmayiermnyihrdlraanilvqdnlvckiadfglarl--iedne-ytarggakfpikwt 432
|||:| | | | | | | | | | | :|:|:|:|:| :|:|:| | :| :| :| :| :| :|
Qy 347 GMAYLESRNYIHRDLAARNVLVGEHNIYKVADFGLARVKVDNEDIYESRHEIKLPVKWT 406

Db 433 apeaalfgkftiksdvwsfgillitelvtkgrvpvgmnnrevleqvergyrmqcpqgcpp 492
||| | | :|:|:|:|:|:|:| | :|:|:|:| | :|:|:| | :| :| :| :| :| :|
Qy 407 APEAIRSNKFSIKSDVWSFGILLYEIIITYGKMPYSGMTGAQVIQMLAQNRYRLPQPSNCPQ 466

Db 493 slhdvmvgcwkrepeertfeylqsfledyftatepqyqpgdn 535
: :|:|:|:| | | | | | | :| | | | | | :| :| :| :| :| :|
Qy 467 QFYNIMLECWNAEPKERPTFETLRWKLEDYF-ETDSSYS 508

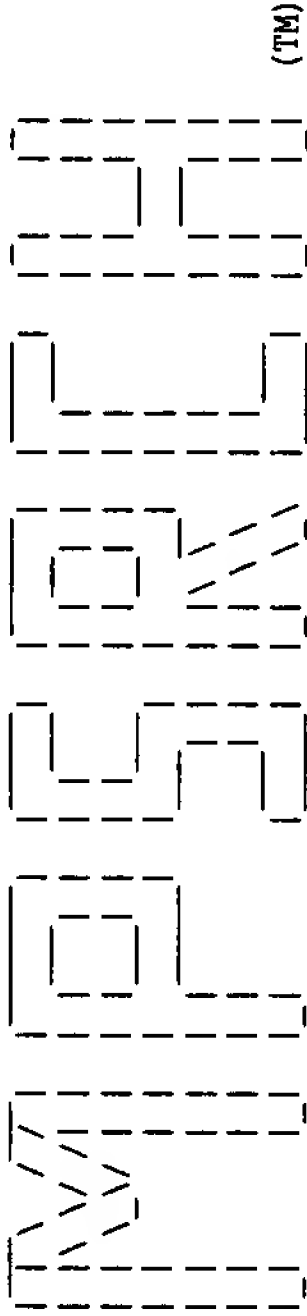
RESULT 5
ENTRY      S29626 #type complete
TITLE      protein kinase yrk (EC 2.7.1.1-) - chicken
ORGANISM   #formal_name Gallus gallus #common_name chicken
DATE       06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
ACCESSIONS S29626; S29553
REFERENCE   S29553
#authors    Sudol, M.; Greulich, H.; Newman, L.; Sarkar, A.; Sukegawa,
              J.; Yamamoto, T.
#submission submitted to the EMBL Data Library, October 1992
#description Novel yes-related kinase, yrk, is expressed at elevated
              levels in neural and hematopoietic tissues.
#accession  S29626
##status     preliminary
##molecule_type mRNA
##residues   1-536 ##label SUD
##cross-references EMBL:X67786

CLASSIFICATION #superfamily protein-tyrosine kinase src; protein kinase
                homology; SH2 homology; SH3 homology
```


May 31 09:27

US-08-426-509-6 frag

1



CC activity.
SQ Sequence 505 AA;

Query Match 97.7%; Score 3734; DB 14; Length 505;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;
Matches 504; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Db 1 msnicqrlewlepylpcclsteadkstvienpgalcsppqsrhghyfvalfdyqartaed 60
|||||
Qy 6 MSNICQRLWEYLEPYLPCCLSTEADKSTVIENPGALCSPQSRHGHYFVAFDYQARTAE 65

Db 61 lsfragdk-lqvltdlhqegwwfarhlekrddgssqqlqgyipsnyvaedrslqaepwffg 119
|||||
Qy 66 LSFragDKKLQVLDTLHEGWWFARHLEKRDRDSSQQLQGYIPSNYVAEDRSLQAEPWFFG 125

Db 120 aigrsdaekqllysenktgsfliresesqgfe slsvldgavvkhyrikrldegffltr 179
|||||
Qy 126 AIGRSDAEKQLLYSENKTGSFLIRESESGKGEFSLSVLDGAVVKHYRIKRLDEGGFFLTR 185

Db 180 rrifstlnefvshytktsdglcvklgkpc lkiqpapfdlsyktvdqweidrnsiqllkr 239
|||||
Qy 186 RRIFSTLNEFVSHYTKTSDGLCVKLKPC LKIQVPAPFDLSYKTVDQWEIDRNSIQLLKR 245

Db 240 lgsqgqgevweglwnttpvavtklpgsmdpndflreaqimknlrhpklqliyavctle 299
|||||
Qy 246 LGSQGQGEVWEGWLWNTTPVAVKTLKPGSMDPNDFLREAQIMKNLRHPKLIQLYAVCTLE 305

Db 300 dpiyiitelmrhgs lqeylqndtqgskihltqqvdmaaqvasgmaylesrnyihrdlaarn 359
|||||
Qy 306 DPIYIITELMRHGS LQEYLQNDTGSKIHLTQQYDMAAQVASGMAYLESRNYIHRDLAARN 365

Db 360 vlvgehniykvadfglarvfkvdnediyesrheiklpvktapeairnskfiks dwsf 419
|||||
Qy 366 VLVGEHNIYKVADFGLARVFKVDNEDIYESRHEIKLPVKWTAPEAIRSNKFSIKSDVWSF 425

Db 420 gillyeitygkmpysgmtgaqv iqlaqnyrlpqpsncpqgfynimlecwnaepkerpt 479
|||||
Qy 426 GILLYEITYGKMPYSGMTGAQVIQMLAQNYRLPQPSNCPQQFYNIMLECWNAEPKERPT 485

Db 480 fetlrwkledyfetdssysdannfir 505
|||||
Qy 486 FETLRWKLEDYFETDSSYSDANNFIR 511

RESULT 2
ID R41941 standard; Protein; 505 AA.
AC R41941;
DT 10-MAR-1994 (first entry)
DE pTK gene LpTK-2 prod.
KW pTK; protein tyrosine kinase; catalytic domain; c-kit; megakaryocyte;
KW lymphocyte; amplification; primer; polymerase chain reaction; PCR.
OS Homo sapiens.
PN W09315201-A.
PD 05-AUG-1993.
PF 22-JAN-1993; U00586.
PR 22-JAN-1992; US-826935.
PA (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
PI Avraham H, Cowley S, Groopman J, Scadden D;
DR WPI; 93-320330/40.
DR N-PSDB; Q49754.
PT New protein tyrosine kinase genes and proteins encoded by genes -
PT are of human mega-karyocytic origin
PS Claim 3; Fig 5; 60pp; English.
CC pTK genes were identified using two sets of degenerative

CC oligonucleotide primers: a first set which amplifies all pTK DNA
CC segments (Q49743-44), and a second set which amplifies highly
CC conserved sequences present in the catalytic domain of the c-kit
CC subgroup of pTKs (Q49745-46). The pTK genes identified are described
CC in Q49747-57 and R41897-02.
CC The lptks are expressed in lymphocytic cells, as well as
CC megakaryocytic cells. The partial and full-length LpTK2 gene
CC sequences are given in Q49749 and Q49754 respectively. The
CC protein sequence corresp. to Q49749 is claimed (claim 7) and
CC stated as given in the specification, however is missing from
CC the publication.
SQ Sequence 505 AA;

Query Match 97.7%; Score 3734; DB 8; Length 505;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;
Matches 504; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Db 1 msnicqrlewlepylpcclsteadkstvienpgalcsppqsrhghyfvalfdyqartaed 60
|||||
Qy 6 MSNICQRLWEYLEPYLPCCLSTEADKSTVIENPGALCSPQSRHGHYFVAFDYQARTAE 65

Db 61 lsfragdk-lqvltdlhqegwwfarhlekrddgssqqlqgyipsnyvaedrslqaepwffg 119
|||||
Qy 66 LSFragDKKLQVLDTLHEGWWFARHLEKRDRDSSQQLQGYIPSNYVAEDRSLQAEPWFFG 125

Db 120 aigrsdaekqllysenktgsfliresesqgfe slsvldgavvkhyrikrldegffltr 179
|||||
Qy 126 AIGRSDAEKQLLYSENKTGSFLIRESESGKGEFSLSVLDGAVVKHYRIKRLDEGGFFLTR 185

Db 180 rrifstlnefvshytktsdglcvklgkpc lkiqpapfdlsyktvdqweidrnsiqllkr 239
|||||
Qy 186 RRIFSTLNEFVSHYTKTSDGLCVKLKPC LKIQVPAPFDLSYKTVDQWEIDRNSIQLLKR 245

Db 240 lgsqgqgevweglwnttpvavtklpgsmdpndflreaqimknlrhpklqliyavctle 299
|||||
Qy 246 LGSQGQGEVWEGWLWNTTPVAVKTLKPGSMDPNDFLREAQIMKNLRHPKLIQLYAVCTLE 305

Db 300 dpiyiitelmrhgs lqeylqndtqgskihltqqvdmaaqvasgmaylesrnyihrdlaarn 359
|||||
Qy 306 DPIYIITELMRHGS LQEYLQNDTGSKIHLTQQYDMAAQVASGMAYLESRNYIHRDLAARN 365

Db 360 vlvgehniykvadfglarvfkvdnediyesrheiklpvktapeairnskfiks dwsf 419
|||||
Qy 366 VLVGEHNIYKVADFGLARVFKVDNEDIYESRHEIKLPVKWTAPEAIRSNKFSIKSDVWSF 425

Db 420 gillyeitygkmpysgmtgaqv iqlaqnyrlpqpsncpqgfynimlecwnaepkerpt 479
|||||
Qy 426 GILLYEITYGKMPYSGMTGAQVIQMLAQNYRLPQPSNCPQQFYNIMLECWNAEPKERPT 485

Db 480 fetlrwkledyfetdssysdannfir 505
|||||
Qy 486 FETLRWKLEDYFETDSSYSDANNFIR 511

RESULT 3
ID R39706 standard; Protein; 536 AA.
AC R39706;
DT 23-DEC-1993 (first entry)
DE Human pp60 c-src protein.
KW Endothelial; tyrosine kinase protein; pp60 c-src; ss.
OS Homo sapien.
PN W09314193-A.
PD 22-JUL-1993.
PF 05-JAN-1993; US00445.

May 31 09:27

US-08-426-589-6.18pg



06-JAN-1992; US-820011.
(UYUA) UNIV YALE.
Bell L, Luthringer DJ, Madri JA, Warren SL;
WPI; 93-243209/30.
P-PSDB; R39705.
Genetically engineered endothelial cells - which exhibit enhanced
cell migration, urokinase-type plasminogen activator activity,
and reduced mononuclear cell adhesion and fibronectin prodn
Disclosure; Page 75-77; 91pp; English.
The DNA encoding a portion or (more preferably) the entire pp60
c-src polypeptide (Given in Q46688) is used to transform endothelial
cells. Transformed cells produce increased amounts of pp60 c-src and
have improved therapeutic properties. They migrate at faster rates
than non-transformed counterparts; have an enhanced ability to
inhibit the formation of thrombi and/or dissolve thrombi once they
have formed and exhibit reduced mononuclear cell adhesion. They can
also be used to improve the success of surgical procedures such as
coronary angioplasty, heart bypass surgery, vessel graft and stent
implantation.
Sequence 536 AA;

Query Match - 43.9%; Score 1679; DB 8; Length 536;
Best Local Similarity 51.3%; Pred. No. 2.18e-143;
Matches 238; Conservative 97; Mismatches 107; Indels 22; Gaps 14;

```
Db      89 fvalydyesrtetdlsfkger-lqivntegdwlahslst---g---q-tgyipsnyv 140
        |||:::|| |::| ::|:: ||::| |::|: | |::|||
QY     52 FVALFYQARTAEDSLFRGDKKLVLDTHLGWFWFARHLEKRDRDSSQQLGYTPSNYV 111
```

```

141 apsdsiqaeewyfgkitreserlllnaenprgtflvresettkgacyclsvsdfdnakgl 200
    | :||| |:}| |::: ||::| |:|:||||: ||: ||| :| | :|
112 AEDRSIQAEPPFFGCAIGRSDAEKOLLYSENKTGSFLIRESEOKGEFSLSV-LDGA-V 167

```

201	nvkhykirkldeggyfyt	rtqfnsllqlvayysk	haglchr	lttvcptsk	pgtqgla	259
	:	:	:	:	:	
168	-VKHVAIKRLDEGGFFL	TRRRIFSTLNEFVSHY	TKTSDGLCVKLGKPC	LIKIQVPAPFDLS	226	
	:	:	:	:	:	

Db 260 -k--daweipreslrlevlkggcfcgvmgtwngttrrvaiktlkpgtmspeafleaqv 316
| | | | |:| |:| | | | | | | | | | | | | | | | | | | |
QY 227 YKTVDWEIDRNSIQLLKRLGSGGFGEVWGLWNNTTPAVKTKPGSMDPNDFLEAAQI 286

317 mkkllrheklvayavvs-eepiyivteymskgsllfdlkgctgkyrlrplqldmaaaias 375
 ||||| ||:||||| : |:|||||:| ||| :|| :|| :|| :|||:|||||
 287 mknlrhpkllioyavctledpilyitelmrhgsgloeylondtfgskihltooydvaaoavas 346

```

376 gmayvermnyvhrdlraanilvgenlvckvadfglarl--iedne-ylarqgakfpikwt 432
      |||||  ||:|||||  |:|||||  : ||||| |||||  : |||||  : |||: |||
347 gmaylesrnythrdlaarnvlvgchntlykvaDEGLARVEKVNED IYESSHEIKIPVKWT 406

```

[illegible]

```

Db      493 slhldmccwrkepeertfeylqafledyftstepayqgenl 536
      : :: :| | | | | | | : | | | | | | | : | :: :|
ov      467 OFYNIMLECNNAEPKERTFEETLRWKIEDYE-ETDSYSDANNF 509

```

RESULT 4
ID R39705 standard: protein: 533 AA

AC R39703;
DT 23-DEC-1993 (first entry)
DE Chicken pp60 c-src protein.
KW Endothelial; tyrosine kinase protein; pp60 c-src; ss.

May 31 09:27

US-08-426-509-6.18g



OS	Gallus gallus.
PN	W09314193-A.
PD	22-JUL-1993.
PP	05-JAN-1993; US00445.
PR	06-JAN-1992; US-820011.
PA	(UYYA) UNIV YALE.
PI	Bell L, Luthringer DJ, Madri JA, Warren SL;
DR	WPI; 93-243209/30.
DR	P-PSDB; R39705.
PT	Genetically engineered endothelial cells - which exhibit enhanced
PT	cell migration, urokinase-type plasminogen activator activity,
PT	and reduced mononuclear cell adhesion and fibronectin prodn
PS	Disclosure; Page 64-66; 91pp; English.
CC	The DNA encoding a portion or (more preferably) the entire pp60
CC	c-src polypeptide (Given in Q46687) is used to transform endothelial
CC	cells. Transformed cells produce increased amounts of pp60 c-src and
CC	have improved therapeutic properties. They migrate at faster rates
CC	than non-transformed counterparts; have an enhanced ability to
CC	inhibit the formation of thrombi and/or dissolve thrombi once they
CC	have formed and exhibit reduced mononuclear cell adhesion. They can
CC	also be used to improve the success of surgical procedures such as
CC	coronary angioplasty, heart bypass surgery, vessel graft and stent
CC	implantation.
SQ	Sequence 533 AA;

Query Match 43.8%; Score 1674; DB 8; Length 533;
Best Local Similarity 51.1%; Pred. No. 6.48e-143;
Matches 237; Conservative 98; Mismatches 107; Indels 22; Gaps 15;

Db	86	fvalydyesrtetdlsfkkgerr-lqivmntegdwla	lsl-t-tg---q-tygpsnyv	137
		: : :	: : :	
QV	52	FVALFDYQAOTAEEDLSFRAGDKKLOVLDTHEGWWFARHLEKKRDGSSOQLGYIPSNYV	111	

```

138 apsdsiqaeewfygkitrreserllnpenprgtflvresettkgayclsvsfidnakgl 197
    | 1 : ||| 1 : || 1 : : 1 : || 1 : ||| : ||| : ||| : ||| :
QV 112 AEDRSLOAEPPFFGAIGRSDAEEKOLLYSENKTSGLIRSESOKGFSLSV--LDGA--V 167

```

```

198 nvkhykirkldsggfyitsrtqfsslqqlvayyskhadglchrltnvcptsk-pqtqgla 256
      |||||:| | |||:| | |||:| | |||:| | |||:| | |||:| | |||:| | |||:
168 -vkhyhikrldeggefflrrrrifstinefshytktsdglcvklgkckikiovpapedls 226

```

```

Db      257  -k--daweipreslrlevklgqcfcgevwmgtwngttrvaiktlkpgnmspeaflqeaqv 313
      + + + + + : + + + + + + + + + + + + + + + + + + + + + + + + + + +
Ov      227  YKTVDWEIDRNSIOIKRIGSGOGEVFEGLWNNTPVAKTKIPGMDPNDEIREAOI 286

```

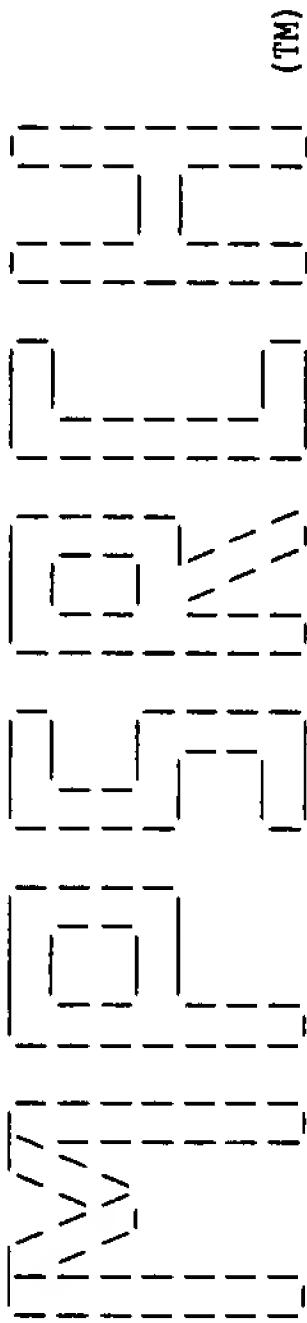
Db 314 mkkllrheklvayavvs-eepiyivteymskgslldfllkgmgkyrlpqlvDMAAqias 372
 ||||| ||:||||| : ||:||||| || ||| : || : ||| |||: ||
 Ov 287 msnlrpklolayavctiedpyiitellmrhgsloeylonDgSKtItTOOYDMAAovvas 346

```
Db      373 gmayvermnyhrdlaanilvgenlvckvadfglarl--iedne-ytarqgakfpikwt 429
        ||||| | | | | | | | | : ||||||| : : : : | : | : | | |
Ov      347 GMAYTFLSRNYTHRDLAARNVIAGEHNTYKVADEGLARVEKYDNEETIYESRHETIKLPVKWT 406
```

Db 430 apeaalgrfrikdsvsflllteltttkgrvpygmrvldqvergyrmpcppece 489
||||| : ||||| ||||| : | : ||| : ||| : ||| : ||| :
Ov 407 APEALPNKKESTIKSDVSEGLIIVEITTYCKMPSYSCMTGCAVIOMIAONVPIPOP SNCPO 466

Db 490 slhdlmcqwrdrpeertfeylqafledyftstetpyqpqgenl 533
 : :: | | : | | | | | : | | | | | : :: | : | :
Ov 467 qevnvmiefcwnafekkedptfeetrdwkienvk-ftntssysdanne 500

RESULT 5
ID R14201 standard; Protein; 417 AA.



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Fri May 31 11:36:53 1996; MasPar time 1686.87 Seconds
1162.552 Million cell updates/sec

Tabular output not generated.

Title: >US-08-426-509-5
Description: (1-2770) from US08426509.seq
Perfect Score: 2770
N.A. Sequence: 1 CCGGACTGGTCGAAAGACAG.....AAGTACCAAAAAAAAAA 2770
Comp: GGCCTGACCAGCTTCTGTC.....TTCATGGTTTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 264399 seqs, 353985056 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embl-new11
1:BCT 2:FUN 3:INV1 4:INV2 5:INV3 6:MAM 7:ORG 8:PLN
9:PRI1 10:PRI2 11:PRI3 12:PRO1 13:PRO2 14:ROD 15:SYN
16:UNC 17:VRT 18:VIR
Database: genbank91
19:BCT1 20:BCT2 21:BCT3 22:BCT4 23:BCT5 24:BCT6 25:BCT7
26:INV1 27:INV2 28:INV3 29:INV4 30:INV5 31:MAM1 32:MAM2
33:PAT1 34:PAT2 35:PAT3 36:PHG 37:PLN1 38:PLN2 39:PLN3
40:PLN4 41:PLN5 42:PLN6 43:PLN7 44:PRI1 45:PRI2 46:PRI3
47:PRI4 48:PRI5 49:PRI6 50:PRI7 51:PRI8 52:PRI9 53:ROD1
54:ROD2 55:ROD3 56:ROD4 57:ROD5 58:ROD6 59:ROD7 60:STR
61:SYN 62:UNA 63:VRL1 64:VRL2 65:VRL3 66:VRL4 67:VRL5
68:VRL6 69:VRT1 70:VRT2 71:VRT3
Database: genbank-new11
72:BCT1 73:BCT2 74:INV1 75:INV2 76:MAM 77:PHG 78:PLN
79:PRI1 80:PRI2 81:PRI3 82:ROD 83:STR 84:SYN 85:UNA
86:VRL 87:VRT
Database: u-emb144_91
88:part1

Statistics: Mean 12.445; Variance 5.270; scale 2.361

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description	Pred. No.
1	2619	94.5	2863	47	HSU00803	Human SRC-like tyrosi	0.00e+00
2	2461	88.8	2810	47	HSU22322	Human nuclear tyrosin	0.00e+00
3	1343	48.5	4590	59	RNU09583	Rattus norvegicus Spr	0.00e+00
4	1148	41.4	2602	54	MMIYKMR	M.musculus mRNA for i	0.00e+00
5	1066	38.5	2028	55	MUSBSK	Mus musculus protein	0.00e+00
6	194	7.0	1829	30	SLSRK1	S.lacustris srk1 mRNA	1.95e-148
7	189	6.8	1702	30	SLSRK4	S.lacustris srk4 mRNA	1.25e-143
8	176	6.4	1253	30	SLSRK2	S.lacustris srk2 mRNA	3.57e-131
9	164	5.9	1190	30	SLSRK3	S.lacustris srk3 mRNA	9.78e-120
10	162	5.8	182	88	RN02888	Rattus norvegicus Spr	7.77e-118
11	127	4.6	1820	71	U01350	Torpedo californica p	6.19e-85
12	121	4.4	2019	29	HYDSTK	H.attenuata src-relat	2.24e-79
13	117	4.2	1518	50	HUMLYNTK	Human Lyn B protein (1.10e-75
14	117	4.2	2298	50	HUMLYN	Human Lyn mRNA encodi	1.10e-75
15	110	4.0	4550	53	MMCYES	M.musculus c-yes mRNA	2.89e-69
16	108	3.9	2200	58	RNFCGR	R.norvegicus FGR mRNA	1.94e-67
17	107	3.9	4517	49	HUMCYES1	Human c-yes-1 mRNA.	1.58e-66
18	105	3.8	1755	71	U01349	Torpedo californica p	1.04e-64
19	106	3.8	1982	70	GGYRKA	G.gallus yrk mRNA.	1.28e-65
20	99	3.6	100	10	HS63F2F	H.sapiens CpG DNA, cl	2.81e-59
21	99	3.6	100	79	HS63F2F	H.sapiens CpG DNA, cl	2.81e-59
22	101	3.6	1804	52	HUMXYES2	human c-yes-2 gene.	4.39e-61
23	100	3.6	1852	71	XLYES	Xenopus laevis mRNA f	3.52e-60
24	98	3.5	100	81	HS57H7R	H.sapiens CpG DNA, cl	2.24e-58
25	98	3.5	100	10	HS57H7R	H.sapiens CpG DNA, cl	2.24e-58
26	98	3.5	2431	71	XLFYNR	Xenopus fyn mRNA enco	2.24e-58
27	98	3.5	2446	71	XELFYNC	X.laevis c-fyn protei	2.24e-58
28	96	3.5	2526	71	XHCFYN	Xiphophorus c-fyn (Xf	1.40e-56
29	97	3.5	3438	58	RATLYNTYRX	Rat lyn protein non-r	1.77e-57
30	93	3.4	1538	58	RATLYNBTYR	Rat lyn B protein tyr	6.83e-54
31	93	3.4	1601	58	RATLYNATYR	Rat lyn A protein tyr	6.83e-54
32	93	3.4	1626	57	MUSSRCPP6	Mouse neuronal proto-	6.83e-54
33	93	3.4	2044	56	MUSLYN	Mouse lyn protein mRN	6.83e-54
34	92	3.3	2133	53	MMCFCGRMR	Murine c-fgr mRNA.	5.34e-53
35	92	3.3	2179	53	MMCFGR	Murine c-fgr53 proto-	5.34e-53
36	89	3.2	2435	52	HUMSLK	Human src-like kinase	2.50e-50
37	90	3.2	2647	49	HUMCSYNA	Human c-syn protoonco	3.23e-51
38	88	3.2	2703	56	MUSLYNB	Mouse lyn B protein t	1.93e-49
39	88	3.2	2766	56	MUSLYNA	Mouse lyn A protein t	1.93e-49
40	86	3.1	102	79	HS63F2R	H.sapiens CpG DNA, cl	1.13e-47
41	86	3.1	102	10	HS63F2R	H.sapiens CpG DNA, cl	1.13e-47
42	85	3.1	1701	70	GGCYES	Chicken mRNA for p61	8.63e-47
43	85	3.1	1805	70	GGYES	G.gallus mRNA for yes	8.63e-47
44	86	3.1	1960	53	MMHCK	Mouse hck gene for ty	1.13e-47
45	86	3.1	2002	55	MUSBMK	Murine macrophage gen	1.13e-47

ALIGNMENTS

RESULT	1	HSU00803	2863 bp	mRNA	PRI	25-MAY-1994
LOCUS		Human SRC-like tyrosine kinase (FRK) mRNA, complete cds.				
DEFINITION		Human SRC-like tyrosine kinase (FRK) mRNA, complete cds.				
ACCESSION		U00803				
KEYWORDS		.				
SOURCE		human.				
ORGANISM		Homo sapiens				
		Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;				
		Eutheria; Primates; Catarrhini; Hominiidae; Homo.				

WKKK3

REFERENCE 1 (bases 1 to 2863)
AUTHORS Lee,J., Wang,Z., Luoh,S.M., Wood,W.I. and Scadden,D.T.
TITLE Cloning of FRK, a novel intracellular SRC-like tyrosine kinase-encoding gene
JOURNAL Gene 138, 247-251 (1994)
MEDLINE 94774047

REFERENCE 2 (bases 1 to 2863)
AUTHORS Scadden,D.T.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1993) D.T. Scadden, New England Deaconess Hospital, Hematology/Oncology, 185 Pilgrim Road, Boston, MA, USA, 02215

COMMENT NCBI gi: 392887
FEATURES Location/Qualifiers
source 1..2863
/organism="Homo sapiens"
/cell_line="BL979"
/cell_type="lymphocytehepatoma"
/tissue type="lymphoid"
448..1965
CDS
/gene="FRK"
/note="NCBI gi: 392888"
/codon_start=1
/product="SRC-like tyrosine kinase"
/translation="MSNICQRLWEYLEPYLPCLISTEADKSTVIENPGALCSPQSQRHG
HYFVALFDYQARTAEIDLFRAGDKLOVLDTLHEGWFFARHLEKRRDGSSQQLQGYIPS
NYVAEDRSLOAEPWFPGAIGRSDAEKQLLYSENKTGSFLIRESESQGFESLSVLDGA
VVKHVRIRKRLDEGGFFLTRRRIFSTINEFVSHYTKTSDGLCVLKGKPCIKIQVPAPFD
LSYKTYDQWEIDRNSIQLKRLSGQGFGEWEGLNWNTTPVAVKTLKP GSMDPNDFLR
EAQIMKNL RHPKLIQLYAVCTLEDPIYIITELMRHGSLOEYLQNDTGSKIHLTQQVDM
AAQVSGMAYLESRNYIHRDLAARNVLGEHNIYKVADFGLARFKVDNEDIYESRHE
IKLPVKWTAPAIRSNKFSIKSDVSWFSFGILLYEIITYGKMPYSGMTGAQVIQMLAQNY
RLQPSNCPQQFYNIMLECWNREPKEPTFETLRWKLEDYFETDSSYSDANNFIR"
BASE COUNT 895 a 550 c 624 g 794 t
ORIGIN

Query Match 94.5%; Score 2619; DB 47; Length 2863;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;
Matches 2666; Conservative 0; Mismatches 5; Indels 7; Gaps 6;

Db 164 gataagaagaaagacaccttcctagtgagcagctgccagctcctgtcagtttgccct 223
|||||
Qy 82 GATAAGAAGAAAGACACCTTCCTAGTGAGCAGCTGCCAGCTCCTGCTCAGTTTGCCT 141
|||||

Db 224 cggggtagcactccagccacagaaaagcggtaagtcctcctccagtaggacttgct 283
|||||

Qy 142 CGGGGTAGCACCTCCAGCCACAGAAAGCAAGCGGTAAGTCTCTCCAGGTAGGACTTGCT 201
|||||

Db 284 gcaaccacagctgctggactgatctgaaacgggactttgcatactctccgaagtatggtga 343
|||||

Qy 202 GCAACCCAGCTGCTGGACTGATCTGAAACGGGACTTTGCATACTCTCCAGGTAGGACTTAAG 261
|||||

Db 344 gttgtgctgacttcaaaagttgcctggtgaaagaagataaagttggatccagagagactaag 403
|||||

Qy 262 GTTGTGCTGACTTCAAAAGTTGCCCTGCTGAAGGAAGATAAAGGTGGATCGCAGAGACTAAG 321
|||||

Db 404 gggagagggagaagccctgctcctcttctccccaccaggcacaatgagacaacatctgtc 463
|||||

Qy 322 GGGAGAGGGAAGCCCTGCTCTCTCTCCCCACCAAGGCACAATGAGCAACATCTGTC 381
|||||

Db 464 agaggtctgggagtagctacctaagaccctatctccccctgtttgtccacggaggcagacaagt 523
|||||

Qy 382 AGAGGCTCTGGGAGTACCTAGAACCCTATCTCCCTGTGTTGTCACGGGCGACAGCAAGT 441
|||||

Db 524 caaccgtgattgaaaatccaggggcccttgcctccccagtcacagaggcatggccact 583
|||||

Qy 442 CAACCGTCATTGAAAATCCAGGGCCCTTGCTCTCCCCAGCTCACAGGGCATGGCCACT 501
|||||

Db 584 actttgtgcttgtttgattaccaggctcggactgctgaggacttgagcttcccgagcag 643
|||||

Qy 502 ACTTTGTGGCTTGTTCATTACCAAGGCTCGGACTGCTGAGGACTTCGAGCTTCGAGCAG 561
|||||

Db 644 gtgacaaaacttcaagttcttgacactttbgcatgagggctggtgttgccagacacttgg 703
|||||

Qy 562 CTGACAAAACTTCAAGTTCTGGACACTTTGCATGAGGGCTGCTGGTTGCCACACACTGG 621
|||||

Db 704 agaaaagacgagatggtccagtcagcaactacaaggtatatctccttctaactacgtgg 763
|||||

Qy 622 AGAAAAGACGAGATGGCTCCAGTCAGCAACTACAAGGCTATATTCTCTTAACCTACGTGG 681
|||||

Db 764 ctgaggacagaagcctacaggcagagccgtggtctcttgaggaatcgggaagatcagatg 823
|||||

Qy 682 CTGAGGACAGAAGCCTACAGGGCAGACCCCTGCTTCTTTGGAGCAATCGGAACATCAGATG 741
|||||

Db 824 cagagaaaactattatatattcagaaaaaacagaccggttctcttaatacagagaaaagtg 883
|||||

Qy 742 CAGAGAAAACAACATTATATTGAGAAAACAAGACCGGTTCTCTTAATCAGAGAAAGTG 801
|||||

Db 884 aaagccaaaaagagaattctctcttcagttttagatgagcagctgtgtaaacactaca 943
|||||

Qy 802 AAAGCCAAAAAGGAGAAATTCCTCTTTTCAGTTTAGATGCGCAGCTGTAAAACACTACA 861
|||||

Db 944 gaattaaaagactggtatgaaggggatttttctcacggaagaagaatcttttcaacac 1003
|||||

Qy 862 GAATTTAAAGACTGGATGAAGGGGATTTTTCACGCGCAAGAACAAATCTTTTCAACAC 921
|||||

Db 1004 tgaacgaatttgtgagccactacaccaagacaagtgaacgcctgtgtcaagctgggga 1063
|||||

Qy 922 TGAACGAATTTGTGAGCCACTACACCAACACAAGTGAACGGCCTGTGTCAAGCTGGGGA 981
|||||

Db 1064 aaccatgcttaagatccaggtccaggtccaggtccatttgattgtcgtataaaaccgtgacc 1123
|||||

Qy 982 AACCATGCTTAAAGATCCAGGTCGCCAGCTCCATTGATTGTGCTATAAACCGTGGACC 1041
|||||

Db 1124 aatggagatagaccgcaactccatcacagcttctgaagcattgggcatcgtggtcagtttg 1183
|||||

Qy 1042 AATGGGAGATAGACCGCAACTCCATACAGCTTCTGAAGCCATTGGGATCTGGTCAGTTTG 1101
|||||

Db 1184 gcgaagtatgggaaggtctgtggaacaataaccactccagtagcagtgaaacattaaaaac 1243
|||||

Qy 1102 GCGAACTATGGGAAGGCTGTGTGAAACAATACCCTCCACTAGCAGTGAAACATTTAAAC 1161
|||||

Db 1244 caggttcaatggatccaaaatgacttcctgagggagcacagataaatgaagaacctaaagac 1303
|||||

Qy 1162 CAGGTTCAATGGATCCAAATGACTTCCTGAGGGGCGCAGAGATAATGAAGAACCTTAAGAC 1221
|||||

Db 1304 atccaaagcttatccagcttttatgctgtttgcaactttaagaagatccaatttatatta 1363
|||||

Qy 1222 ATCCAAAGCTTATCCAGCTTTATGCTGTGTTGCACTTTAGAAGATCCAATTTATATTATA 1281
|||||

Db 1364 cagagttgatgagacatggaagtctgcaagaatatctccaaaaatgacactggatcaaaaa 1423
|||||

Qy 1282 CAGAGTTGATCAGACATGGAAGCTCTGCAAGAAATATCTCCAAAATGACACTGGATCAAAA 1341
|||||

Db 1424 tccatctgactcaacaggtagacatggcggcagagttgcctctggaatggcctatctgg 1483
|||||

Qy 1342 TCCATCTGACTCAACAGGTAGACATGGCGGCACAGGTTGCCTCTCGAATGSCCTATCTGG 1401
|||||

Db 1484 agtctcggaactacattcacagagatctgggctgccagaaatgtcctcgttgggtgaacata 1543
|||||
Qy 1402 AGTCTCGCAACTACATTTCACAGACATCTGGCTGCCAGAAATGTCCTCGTTGGTGAACATA 1461

Db 1544 atatctacaaagtagcagattttggacttggcagagttttttaagglagataaatgaagaca 1603
|||||
Qy 1462 ATATCTACAAAGTAGCAGATTTTGGACTTGGCAGAGTTTTTAAGGTAGATAAATGAAGACA 1521

Db 1604 tctatgaaactagacacgaaataaagctgccggtgaagtggactgcgcgcgaagccattc 1663
|||||
Qy 1522 TCTATGAATCTAGACACACGAAATAAAGCTGCCGTTGAAGTGGACTGCGCCGGAAGCCATTC 1581

Db 1664 gtagtaataaaattcagcattaaagtcgcgatgtatgggtcaatttggaaatccttctttatgaaa 1723
|||||
Qy 1582 GTAGTAATAAANTCAGCATTAAGTCCGATGTATGGTCAFTTGGAAATCCTTCTTTATGAAA 1641

Db 1724 tcattacttatggcacaatgacctacagtggtatgacaggtgccagggtaaatccagatgt 1783
|||||
Qy 1642 TCATTACTTATPGCCAAAATGCCTTACAGTGGTATGACAGGTGCCCAGGTAATCCAGATGT 1701

Db 1784 tgggtcaaaactatagacttcgcgaaccatccaactgtccacagcaaatittacaacatca 1843
|||||
Qy 1702 TGGCTCAAAACTATAGACTTCGCGCAACCATCCAAGTGTCCACACGCAATTTTACACATCA 1761

Db 1844 tgttgagagtgtggaatgcagagcctaaggaacgacctacatttggacacactgcggttggg 1903
|||||
Qy 1762 TGTTGGAGTGTGGAATGCAGAGCCTTAAGGAACGACCTACATTTTGAGACACTGCGTGGG 1821

Db 1904 aacttgaagactattttgaaacagagactcttcattcagatgcaaaataaacttcaataagat 1963
|||||
Qy 1822 AACTTGAAGACTATTTTGAACAGAGACTCTTCATATTCAGATGCAAAATAACTTTCATAAGAT 1881

Db 1964 gaacactggagaagaatatcaaaataataaagtagcaaaaacaaattcaataataatccattcc 2023
|||||
Qy 1882 GAACACTGGAGAAGAATATCAAAATAATAAAGTAGCAAAACAAATTCMAATAATCCATTCC 1941

Db 2024 aaaatacaatggtatcaaccaactgcacaatcagtttatcctgacatatccaaagtgatag 2083
|||||
Qy 1942 AAAATACAAATGTTATCAACCAACTGCACAATCAGTTTATCCTGCACATATTCAAAGTGATG 2001

Db 2084 gataaagttggccatgtattatgaaaaagattatttgtgcattttattgactgggcaaca 2143
|||||
Qy 2002 GATAAAGTTGGCCCATGTATTATGAAAAAGATTATTTGTGCAATTTATTGACTGGGCAACA 2061

Db 2144 ctgcaggacagtcaaggtgatataataatttccctcactgcctgggtaaaattaagcacacta 2203
|||||
Qy 2062 CTGCAGGACAGTCAAGGTCAATATAATAATTGC-TCACCTGCCTGG-AAAAATTAAGCACACTA 2119

Db 2204 aaccaagttattttctttttaagagatacttaccatttccatttattgtttgaaatgtcg 2263
|||||
Qy 2120 AACCAAGTTATTTTCTTTTAAAGAGATACTTACATTTCCATTTATTGTTTGAATGTCTG 2179

Db 2264 ---atcaagagaatcaacagatgatagtcctcaatttttactcagtgactg-t-tgtagcatt 2319
|||||
Qy 2180 CGATCAAGAGAAATCAACAGATGATAGTCCAAATTTTACTCAGTGATGACTGTGTAGCAATT 2239

Db 2320 ttccctgtttactgattagatggttattcattattcctcagattgctgaatcccatcagg 2379
|||||
Qy 2240 TTCCCTGTTTACTGATTAGAGTGGTTAATTCATTTATTCCTCAGATGTCTGAATCCCATCAGG 2299

Db 2380 ctgttattatgaaggaatttgattgctttgtgcacagcaggacctgtgcttttgagattt 2439
|||||
Qy 2300 CTGTTATTATGAAGGAATTTGATTGCTTTGCTGCACAGCAGGACCTGTGCTTTGAGATTT 2359

Db 2440 ttttttctcttttaaaatatcctgtaactacaatgatgtgtaaagccatgtttaaagtactt 2499
|||||
Qy 2360 TTTTCTCTCTTTTAAATAATCCTGTAACTACAATCATGTGTAAAGCCATGTTAAATGACTT 2419

Db 2500 gattgtacttgagtaaattgcacatttttttctctatgcataaaaaaatgatgcagctgttg 2559
|||||
Qy 2420 GATTGTACTTGCAGTAATTGCACATTTTTTCTATGCATAAAAAAATGATGCAGCTGTG 2479

Db 2560 agaaaacgaagtcttttccattttgcagaaggaaatgatggaatttttctgtacttcaagt 2619
|||||
Qy 2480 AGAAAACGGAAGTCTTTTTCATTTTGCAAGAGGAAATGATGGAATTTTCTCTACTTCAGT 2539

Db 2620 atgtgtcaactgagagtcatacatcattagttttaaactcttaattgagaatcaggttg 2679
|||||
Qy 2540 ATGTGTCAACTGAGAGTCATATACATTACTTTTAAATCTCTTAATATTGAGAATCAGGTTG 2599

Db 2680 caaaacggatgagttattatctatggaatgtgagaaatgtctaataagcccataaaagtct 2739
|||||
Qy 2600 CAAAACGGATGAGTTATTATCTATGGAATGTGAGAAATGTCTAATAGCCCATAAAGTCT 2659

Db 2740 gagaaataggtatcaaaaatagtttaggaaaaatgagaggagaacagtag-attgctgtggc 2798
|||||
Qy 2660 GAGAAATAGGTATCAAAATAGTTTAGGAAAATGAGAGGAGCAACACTAGGATTGCTGTGGC 2719

Db 2799 ctagacttctgagtaattaataaagaaaaagaagtacc 2836
|||||
Qy 2720 CTAGACTTCTCAGTAATTAATAAAGAAAAAGAACTACC 2757

RESULT 2
LOCUS HSU22322 2810 bp mRNA PRI 27-MAR-1995
DEFINITION Human nuclear tyrosine protein kinase Rak mRNA, complete cds.
ACCESSION U22322
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; eukaryote crown group;
Metazoa/Eumycota group; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2810)
AUTHORS Cance,W.G., Craven,R.J., Bergman,M., Xu,L., Alitalo,K. and Liu,E.T.
TITLE Rak, a novel nuclear tyrosine kinase expressed in epithelial cells
JOURNAL Cell Growth Differ. 5 (12), 1347-1355 (1994)
MEDLINE 95210168
REFERENCE 2 (sites)
AUTHORS Cance,W.G., Craven,R.J., Weiner,T.M. and Liu,E.T.
TITLE Novel protein kinases expressed in human breast cancer
JOURNAL Int. J. Cancer 54 (4), 571-577 (1993)
MEDLINE 93293373
REFERENCE 3 (bases 1 to 2810)
AUTHORS Cance,W.G.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-1995) William G. Cance, Surgery Dept., University
of North Carolina, Chapel Hill, NC 27599, USA
NCBI gi: 732527
FEATURES Location/Qualifiers
source 1..2810
/organism="Homo sapiens"
/cell_line="BT-20 breast cancer cell line"
410..1927
/note="nuclear tyrosine protein kinase; Method:
conceptual translation supplied by author. NCBI gi:

732528"
/codon_start=1
/product="Rak"
/translation="MSNICQRLWEYLEPILPCLSTEADKSVIENPGALCSPQSQRHG
HYFVALFDYQARTAEIDLFRAGDKLQVLDTLHEGWFAHLEKRRDSSQQLQGYIPS
NYVAEDRSLOAEAWFFGAIRRSDAEKQLLYSENKTSFGLIRESESKQGEFSILVDGA
VVKHYIRKRLDEGGFFLTRRRIFSTLNEFVSHYTKTSDGLCVKLGKPCIKIQVPAPFD
LSYKTYVDQWEIDRNSIQILKRLSGQFGEVWEGLMNNTTPVAVKTLKPGSMDPNDFLR
EAQIMKNLHRPKLIQLYAVCTLEDPIYIITELMRHCSLOEYLQNDTFSKIHILTQQVDM
AAQVASGMAYLESRNYIHRDLAARNVLGGEHNIYKVADFGLARVFKVDNEDIYESRHE
IKLPVKWTAPEAIRSNKFSIKSDVWSFGILLYEITYGKMPYSGMTGAQVIQMLAQN
RLPQPSNCPQQFYNIMLECNWAEPKERTFETLRWKLEDYFETDSSYSYDANNEIR"
910..952
misc_feature
polyA_site
/note="putative protein nuclear localization sequence"
2810
/note="16 A nucleotides"
BASE COUNT866 a554 c612 g778 t
ORIGIN

Query Match88.8%; Score 2461; DB 47; Length 2810;
Best Local Similarity98.6%; Pred. No. 0.00e+00;
Matches 2612; Conservative0; Mismatches 13; Indels 23; Gaps 20;

Db166cccagctcctgcagcttggcctcggggttagcaccttccagccacacagaaacg-aagccg224
|||||
Qy118CCCAGCTCCTGCTCAGTTTGCCTCGGGGT-AGACCT--CCAGCCACAGAAAGCAAGCCG175

Db225gta-gtctctc-ag-tag-acttggtgcaacccagctgctggactgatctgaaacgg-ac279
|||
Qy176GTAAGTCTCTCCAGGTAGGACTTGCTGCAACCCAGCTGCTGGACTGATCTCAAAACGGGAC235

Db280tttgatactctccgaagtatggtgagttggtgactcaaatccagttgcctggtgaacca339
|||||
Qy236TTTGCTACTCTCCGAAGTATGGTGAGTTGGTGCTGACTTCAAAGTTGCCTGGTGAAGGA295

Db340agataaggtggatcgagagactatgtagagggagaggagagccctgctcctcttctcccca399
|||||
Qy296AGATAAGGTGGATCGCAGAGACTAAGGGGAGAGGGAGAAGCCCTGCTCCTCTTCTCCCCA355

Db400ccaaaggcacaaatgagcaacatctgtcagaggctctggagtagacctagaaacctatctccc459
|||||
Qy356CCAAGGCACAATGAGCAACATCTGTCAGAGGCTCTGGAGTAGCTAGAACCTATCTCCC415

Db460ctgtttgtccacggaggcagacaaagtcacaccgtgattgaaaaatccaggggcccttgctc519
|||||
Qy416CTGTTTGTCCACGGAGGCAGACAAAGTCAACCCGTGATTGAAAATCCAGGGGCCCTTTGCTC475

Db520tcccagtcacagaggcatggccactactttgtggcttggttgattaccaggctcggac579
|||||
Qy476TCCCAGTCACAGAGGCATGGCCACTACTTTGTGGCTTTGTTGATTACCAGGCTCGGAC535

Db580tgctgaggacttgagcttcgagagcaggtgacaaaactcaagttctggacactttgcatga639
|||||
Qy536TGC TGAGGACTTGAGCTTCCAGCAGGTGACAAACTTCAAAGTTCTGGACACTTTGCATGA595

Db640gggctgggtgttggcagacacttgagagaaaagacgagatggctccagtcagcaactaca699
|||||
Qy596GGGCTGGTGTGGCAGACACTTGGAGAAAAGACCGAGATGGCTCCAGTCAGCAACTACA655

Db700aggctatatctcttaactacgtggctgaggacagagcctacagaggagggcgtgtt759
|||||
Qy656AGGCTATATTCCTTCTAACTACGTGGCTGAGCAGAGAAGCCCTACAGGGCAGAGCCGTGGTT715

Db760ctttggagcaatcagaagatcagatgcagagagaaacaactattatattcagaaaaacaagac819

Qy716CTTTGGAGCAATCGGAAGATCAGATGCAGAGAAACAACATATATATTCAGAAAAACAAGAC775

Db820cggttccttctaataatcagagaaagtgaagccaaaaaggagaattctctcttcagtttt879
|||||
Qy776CGGTTCCTTTCTAATCAGAGAAAGTGAACCCAAAAGGAGAAATTCTCTCTTTTCAGTTTT835

Db880agatggagcagttgtaaacactacagaataaaagactggaaggggggatttttct939
|||||
Qy836AGATGGAGCAGTTGTAAACACTACAGAAATTAACAGACTCGATGAAGGGGGATTTTTTCT895

Db940cacggaagaagaatcttttcaacactgaacgaatttgtgagccactacaccaagacaag999
|||||
Qy896CACGGGAAGAATCTTTTCAACACTGAACGAATTTGTGAGCCACTACACCAAGACAAG955

Db1000tgacggcctgtgtcaagctggggaaccatgcttaaatccaggtcccagctccatt1059
|||||
Qy956TGACGGCCTGTGTCAAGCTGGGAAACCATGCTTAAAGATCCAGGTCCCAGCTCCAATT1015

Db1060tgattgtcgtataaaacctggaccaatgggagatgagccgaactccatacagcttct1119
|||||
Qy1016TGATTGCTCGTATAAACCCTGGACCAATGGGAGATAGACCGCAACTCCATACAGCTTCT1075

Db1120gaagcgattgggatctggtcagtttggcgaagtgtgggaaggtctgtggaacaataccac1179
|||||
Qy1076GAAGCGATTGGGATCTGGTCAGTTTGGCGAAGTATGGGAAGGTCTGTGGAACAATACCAC1135

Db1180tccagtagcagtgaacacattaaaaccaggttccaatggatccaaatgacttcctcgaggga1239
|||||
Qy1136TCCAGTAGCAGTCAAAACATTAAAACCAGGTTCAATGGATCCAAATGACTTCCTTGAGGGA1195

Db1240ggcacagataatgaagaacctaaagacatccaaagcttatccagctttatgctgtttgcac1299
|||||
Qy1196GGCACAGATAATGAAGAACCTAAGACATCCAAAGCTTATCCAGCTTATGCTGTTTGCAC1255

Db1300tttagaagatccaatttatattacagagttgatgagacatggaagctgcaagaata1359
|||||
Qy1256TTTAGAAGATCCAAATTTATATTATTACAGTTGATGACACATGGAAGTCTGCAAGATA1315

Db1360tctccaaaatgacactggatcaaaaatccatctgactcaacaggtagacatggcggcac1419
|||||
Qy1316TCTCCAAAATGACACTGGATCAAAAATCCATCTGACTCAACAGGTAGACATGGCGGCACA1375

Db1420ggttgcctctggaatggcctatctggaqtctcggaactacattcacagagatctggctgc1479
|||||
Qy1376GGTTGCCTCTGGAATGGCCTATCTGGAGTCTCGGAACATACATTACAGAGATCTGGCTGC1435

Db1480cagaaatgtcctcgttggtgaacataatatctacaaaagtagcagattttggacttgccag1539
|||||
Qy1436CAGAAATGTCCTCGTTGGTGAACATAATATCTACAAAAGTAGCAGATTTTGGACTTGGCAG1495

Db1540agttttaaggtagataatgaagacatctatgaatcaagacacgaaaaaagctaccggt1599
|||||
Qy1496AGTTTTTAAGGTAGATAATGAAGACATCTATGAATCTAGACACCCGAAATAAGCTGCCGGT1555

Db1600gaagtggactgcgccgaagccattcgtagtaataaattcagcattaagtccgatgatg1659
|||||
Qy1556GAAGTGGACTGCGCCCGAAGCCATTCTAGTAATAAATTCAGCATTAAGTCCGATGATG1615

Db1660gtcatttggaaatccttctttatgaaatcattacttatggcaaaatgccttacagtggat1719
|||||
Qy1616GTCATTGGAAATCCTTCTTTATGAAATCATTACTIONTATGGCAAAATGCCTTACAGTGGTAT1675

Db1720gacaggtgccaggtaatccagatgttggctcaaaactatagacttccgcaaccatccaa1779

Qy 1676 GACAGGTCGCCAGGTAATCCAGATGTTGGCTCAAAACTATAGACTTCCGCAACCATCAA 1735
Db 1780 ctgtccacagacaattttacaacatcatgttgagtgctggaatgcagagcctaaggaacg 1839
Qy 1736 CTGTCCACAGCAATTTTACAACATCATGTTGGAGTCTGGAATGCAGAGCCTAAGGAACG 1795
Db 1840 acctacatttgagacactgcgttggaaccttgaaagactattttgaaacagactcttcata 1899
Qy 1796 ACCTACATTTGAGACACTGCGTTGGAACCTTGAAGACTATTTTGAACAGACTCTTCATA 1855
Db 1900 ttcagatgcaaatcaacttcataagatgaacactggagaagaatatcaaatataaagtag 1959
Qy 1856 TTCAGATGCAAAATAACTTCATAAGATGAACACTGCAGAGAATAATCAAAATAAAGTAG 1915
Db 1960 caaaacaaattcaaaataaatccattccaaatacaaatgttatcaaccaactgcacaat 2019
Qy 1916 CAAAACAAATTCAAATAAT---CCATTCCAAATACAAATGTATCAACCAACTGCACAAT 1972
Db 2020 cagtttatccctgacatatattcaagtgataggataaagttggccatgtattatgaaaaagat 2079
Qy 1973 CAGTTTATCTGCACATATTCAAGTGATAGATAAAGTTGSCCATGCTATTATGAAAAAGAT 2032
Db 2080 tatttctgcattttattgactgggcaacactgcaggacagtcaagtgatataataattc 2139
Qy 2033 TATTTCGCAITTTATTGACTGGCAACACTGCAGGACACTCAAGTCCATATATAATTGC 2092
Db 2140 ctcactgcctggtaaaattaagcacactaaacceagttattttttttaagagatact 2199
Qy 2093 -TCACTGCCTGG-AAAATTAGCACACTAAACCAAGTTATTTTCTTTTAAAGAGATACT 2150
Db 2200 tacatttccatttattgtttgaaatgtcg--atcaagagaatcaacagatgatagccaa 2257
Qy 2151 TACATTCCAITTATTGTTTGAATGTCCGATCAAGAGAATCAACAGATCATAGTCCAA 2210
Db 2258 tttttactcagtgaactg-t-tgtagcattttctctgtttactgattgattgagtggtat-cat 2314
Qy 2211 TTTTACTCACTGATGACTGTCTFAGCATTTTCCTGTTTACTGATTAGAGTGGTTATTCAT 2270
Db 2315 tattctcagattgctgaatcccatcagcgtgttattatgaaggaatttgattgctttgc 2374
Qy 2271 TATTCTCAGATTGCTGAATCCCATCAGCTGTATTATGAAGGAATTTGATTGCTTTGC 2330
Db 2375 tgcaacagcaggacctgctttgagatttttttctctcttttaaaatatcctgtactac 2434
Qy 2331 TGCACAGCAGCCTGCTGCTTTGAGATTTTTCCTCTTTTAAATAATCCTGTAACTAC 2390
Db 2435 aatgatggtaaagccatgttaaatgacttgattgacttgagtaattgcacatttttt 2494
Qy 2391 AATGATGGTAAAGCCATGTTAAATGACTTGATTGCTACTTGGAGTAATTGCACATTTTTT 2450
Db 2495 tctatgcataaaaaaatgatgcagctgttgagaaaaacgaagcttttttcattttgcagaa 2554
Qy 2451 -CTATGCATAAAAAAATGATGCAGCTGTTGAGAAAACGAAGTCTTTTCATTTTTCAGAA 2509
Db 2555 ggaatgatggaatttttctgacttcaagtatgtgtcaactgagatgcataacattagt 2614
Qy 2510 GGAATGATGCAATTTTCTCTACTTCAGTATCTGTCAACTGACAGTCATATACATTAGT 2569
Db 2615 tttaatctcttaattatgagaatcaggttgcaaa-cggatgagttattatctatg-aaat 2672
Qy 2570 TTTAATCTCTTAATATTGAGAATCAGGTTGCAAAACGGATGAGTTATTATCTATGGAAT 2629
Db 2673 gtgagaaatgtctaataagccccataa-gtctgagaaataggtatcaaaaatagtttaggaaa 2731

Qy 2630 CTCAGAAATCTTAATAGCCCAATAAAGCTCTGAGAAATAGGTATCAAAATAGTTTAGGAAA 2689
Db 2732 atgagaggaacagtagtattgtgtgacctagacttctg-gtaattaataaagaaaaa 2790
Qy 2690 ATCAGAGCAGAACAGTAGCATTCGCTCTGGCTAGACTTCTCAGTAATTAATAAGAAAAA 2749
Db 2791 gaagtacc 2798
Qy 2750 GAAGTACC 2757
RESULT 3
LOCUS RNU09583 4590 bp mRNA ROD 09-AUG-1995
DEFINITION Rattus norvegicus Sprague-Dawley src related tyrosine kinase mRNA,
complete cds.
ACCESSION U09583 U02888
KEYWORDS .
SOURCE rat.
ORGANISM Rattus norvegicus
Eukaryotae; Chordata; Eukaryote crown group; Metazoa/Eumycota
group; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia;
Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii;
Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Glires;
Rodentia; Sciurognathi; Myomorpha; Muridae; Rattus.
REFERENCE 1 (bases 1 to 4590)
AUTHORS Avigan,M.I. and Sunitha,I.
TITLE Characterization of GASK, a novel src related tyrosine kinase
JOURNAL Unpublished
REFERENCE 2 (bases 1510 to 1691)
AUTHORS Sunitha,I. and Avigan,M.I.
TITLE A newly identified tyrosine kinase is preferentially expressed in
the gastrointestinal tract
JOURNAL Biochim. Biophys. Acta (1994) In press
REFERENCE 3 (bases 1 to 4590)
AUTHORS Avigan,M.M.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-1994) Mark M. Avigan, Pathology, Georgetown
University School of Medicine, 3900 Reservoir Road N.W.,
Washington, D.C. 20007, USA
COMMENT NCBI gi: 939624
FEATURES Location/Qualifiers
source 1..4590
/clone="GASK"
/clone_lib="RIntTKBS"
/strain="Sprague-Dawley"
/organism="Rattus Norvegicus"
/sex="male"
/tissue_type="small intestine"
/dev_stage="infant"
445..1965
/note="NCBI gi: 939625"
/codon_start=1
/product="src related tyrosine kinase"
/translation="MGSVCMRLWAYLQPFLLPCWSQEADKSWIENPGAYCPCPEANRSQ
GQYFVALFDYEARTAEDLSFHAGDKQLVDLTSHEGWLARHLEKRGPGLGQQLQGYIP
SNYYAEDRSLQAEPWFFGAIKRADAERKQLLYSENQTGAFLIRESQKGFSLSVLDE
GVVKHYRIRLDEGGFFLTRKTFSTLINEFVNYTTSDGLCVKLEKPCIKIQVPTFF
DLSYKTVDQWEIDRNSIQLKRLGSGQFGEVWEGIWNNTTPVAVKTIKPGSMDPNDFL
REAQIMKSLRHPKLIQLYAVCTLEDPIYIITELMRHGS LQEYLDNDGGSKIRLTQQVD
MAAQVASGMAYLESQNYIHRDLAARNVLVGEHNIYKADFGLARVFKVDNEDIYESKH
EIKLPVKWTAPEAIRTNKFSIKSDVWSFGILLYEIITYGKMPYSGMTGAQVIHMLGQN
YRLQPSNCPPEQFYSIMMECNVPEPKQRPTFTLHWKLEDYFEPDSSYSYDTNNFIN"

BASE COUNT 1370 a 937 c 999 g 1284 t

ORIGIN

Query Match 48.5%; Score 1343; DB 59; Length 4590;
Best Local Similarity 82.2%; Pred. No. 0.00e+00;
Matches 1958; Conservative 0; Mismatches 387; Indels 38; Gaps 22;

Db	441	caccatgggcagcgtctgtatgagggctttgggcataccttcagcccttttcccccgtcg	500
Qy	362	CACAATGAGCAACATCTGTCAGAGGCTCTGGGAGTACCTAGAACCCCTATCTCCCCCTGTTT	421
Db	501	gtcccaagaagcagacaaatcagtggtgattgagaatccaggggcctactgtcctccaga	560
Qy	422	GTCCACGGAGGCACACAAGTCAACCGTGATTGAAAAATCCAGGGGCCCTTTGCTCTCCCCA	481
Db	561	ggccaatcgaagcccaaggccagttatttggctgtctgtttgattacgaagctcgtactgc	620
Qy	482	GTCAACA--GAGGC-ATGSCCCTACTTTTGTGGCTTTGTTTGTATTACCAAGGCTCGGACTGC	538
Db	621	agaggacctgagcttccatgcgcggcgacaagctccaagctcttggaacattctcatgaggg	680
Qy	539	TGAGGACTTGACCTTCCGAGCAGGTGACAAACTTCAAGTTCTGGACACTTTGCATGAGGG	598
Db	681	ctgggtggtggccagacatttggaaaagaagggacccggcttaggtcagcagctgcaggg	740
Qy	599	CTGGTGGTTTGGCAGACACTTGGAGAAAAAGACGAGATGGCTCCAGTCAGCAACTACAAGG	658
Db	741	ctacattccttccaattacgtggcagagaccggaggtctccaggcagaaacogtggtttt	800
Qy	659	CTATATTCTTCTTAACGTGGCTGAGGACAGAAAGCTTACAGGCAGAGCGGTGGTTCTT	718
Db	801	cggagcaatcaaaagagcagatgcagaaaagcaacttctgtattccagaaaaccagacggg	860
Qy	719	TGGAGCAATCGGAAGATCAGATGCAGAGAAACAACACTATTATATTACAGAAAAACAAGACCGG	778
Db	861	cgccttctaatcagagagtgagagcagaaaggggtgacttctccctctcagttttaga	920
Qy	779	TTCCCTTTCTAATCAGACAAAGTGAAAGCCAAAAGGAGCAATTCCTCTTTTCAGTTTTAGA	838
Db	921	tgaagtggtgtgaacacactacaggataagaaggttggatgaagtgaggttcttctctcac	980
Qy	839	TGGACGAGTTGTAAACACTACAGAAATTAAGAAGACTGGATGGATGAAGGGGATTTTTCCTCAC	898
Db	981	caggagaaaaacttttcaaccttgaatgaatttgtgaactactacaccacaacaagtga	1040
Qy	899	CGGAAGAAGAAATCTTTTCAACACTGACCGAAATTTGTGAGCCACTACACCAAGACAAGTGA	958
Db	1041	tgggctgtgtcagctggagaagccatgcttaaaagattcaggtaccgaccccttttga	1100
Qy	959	CGGCCTGTGTCAAGCTGGGNAACCAATGCTTAAAGATCCAGGTCCCAGCTCCATTGGA	1018
Db	1101	tttgtcatataaaactgtggaccagtgggagatagaccgcaactccatacagcttctcaa	1160
Qy	1019	TTTGTGCTATAAAACCGTGGACCAATGGGAGATAGACCCGCAACTCCATACAGCTTCTGAA	1078
Db	1161	gcgactggggtctgtgcagtttggagaagtttgggaaggtctgtggaaataataaccactcc	1220
Qy	1079	GCGATTGGGATCTGGTCAGTTTGGCGGAAGTATGGGAAGGTCTGTGGAAACAATACCACCTC	1138
Db	1221	agtggtgtgaaaacgttaaaaccaggttcaatggatccaaaatgacttccctcaggaggcc	1280
Qy	1139	AGTAGCAGTGAAAACATTAAACCAGGTTCAATGGATCCAAATGACTTCCCTGAGGGAGGC	1198
Db	1281	acagataatgaagagcctaagacaccccaagctcatccagctctatgctgtttgcacatt	1340

Qy	1199	ACAGATAATGAAGAACCTTAAGACATCCAAAGCTTATCCAGCTTTATGCTGTTGCACCTT	1258
Db	1341	agaagatccaatttatatttacagaattgatgagacatgggaagctctgcaagaatatct	1400
Qy	1259	ACAAGATCCAATTTATATTATTACAGAGTTGATGAGACATGGAAGTCTGCAAGAAATATCT	1318
Db	1401	ccaaaaatgatgggtgcgaaaaatccgttggactcaacaggtggacatggctgcacaggt	1460
Qy	1319	CCAAAATGACACTGGATCGATCAAAAATCCAATCTGCTCAACAGGTAGACATGGCGGCACAGGT	1378
Db	1461	agcttctggaaatggccctatttggagttccagaaactatatccacagagatctggctgcaag	1520
Qy	1379	TGCCTCTGGAATGGCCTATCTGGAGTCTCGGAACTACATTCAACAGAGATCTGGCTGCCAG	1438
Db	1521	aaatgtccttgggtgaacataatatctataaagtagcagattttggacttgcgaaggt	1580
Qy	1439	AAATGTCTCTGTTGGTGAACATAAATATCTACAAAAGTAGCAGATTTTGGACTTGGCAGAGT	1498
Db	1581	ttttaaggtggataaatgaagacatctatgaatctaaacacagaaaataaagctgccagtga	1640
Qy	1499	TTTTAAGGTAGATAATGAAGACATCTATGAAATCTAGACACGGAATAAAGCTGCCGGTGAA	1558
Db	1641	gtggactgcacccgaagccattcgtactaataaattcagcattaaagtctgatgtgtggtc	1700
Qy	1559	GTGGACTGCCGCCGAGGCCATTGCGTAGTAATAAATTCAGCAATTAAGTCCGATGTATGGTC	1618
Db	1701	ttttggaatcctgctctatgaaatcattacttatggcaaaaatgccttacagcgggatgac	1760
Qy	1619	ATTTCGAAATCCCTCTTTTATGAAATCAATTAATCTATGGCAAAATGCCTTACAGTGGTATGAC	1678
Db	1761	aggtgctcaagttaattcatatgttgggtcaaaactacagactaccacagccatctaactg	1820
Qy	1679	AGGTGCCCAGGTAATCCAGATGTTGGCTCAAAACTATAGACTTCCGCAACCAATCCAACCTG	1738
Db	1821	tccagagcaattctatagcatcatgatggaatgctggaatgtagagcccaagcaacgacc	1880
Qy	1739	TCCACAGCAATTTTACAACTCATGTTGGAGTGTGGAATGCAGAGCCTTAAGGAACGACC	1798
Db	1881	aacgttcgagaccctgcactggaaaacttgaagactactttgaaccagactcttcctattc	1940
Qy	1799	TACATTTGAGACACTGCGTTCGAAACTTGAAGACTATTTTGAACAGACACTCTTCATATTC	1858
Db	1941	tgatacaaaacacttccataaactaaacaacaggaaaaaagaagaataataaattaaagt	2000
Qy	1859	AGATGCAAAATAACTTTCATAAGATGAACACTGGAGAGAAGAAATATCAAAATAAAGTAGCAA	1918
Db	2001	gaaaacgcggaactcaccagtcagaaaaatatcgtcttcaaccaactgtgaaatcagct	2060
Qy	1919	AACAAATTCAAATAATCCATTTCCA-AAAATACAATGTTATCAACCAACTGCACAATCAGTT	1977
Db	2061	tattttgacatattcaggtggtaggatgaactcaaccatgcattctgtatcaaaaaagatc	2120
Qy	1978	TATCCTGCACATATTCAAGTGATAGGATAAAAGTTGGCCCATGTATTATG-----AAAAAGATT	2033
Db	2121	atctgtgtattctattgtagtagaaacgctgcaggacagatataagatgatgtgtcatttt	2180
Qy	2034	ATTTGTGCATTTTATTGACTGGGCCAACACTGCAGGACAGTC-AAGGTCAATATATAATTG-	2091
Db	2181	tccctcactgcctggtaaaatcaatcacactaaatgaagttatttttcttctcttttaaga	2240
Qy	2092	---CTCAGCTGCCCTGG-AAAATTAAGCACACTAAACCAAGTTATTTTCTT---TTTAAAGA	2144
Db	2241	actcctcacatttctatttattgtttgaaatg-c-cgatcaagagaataacacagatgacg	2298

```
Qy 2145 GATACTTACATTTCACATTATTGTTGTAATGTCGGCATCAAGAGAAATCAACAGATGATA 2204

Db 2299 gtg-----tttactctgttactgttgtg--gcattttcctgtttactgattagactggtt 2352
|| |||||| || | ||| ||||||||| ||||||||| ||||||||| ||||||

Qy 2205 GTCCAAFTTTTACTCAGTGATGACTGTGTAGCAATTTTCCTGTTTACTGATTAGACTGGTT 2264

Db 2353 attcattattcctcaaatgctgaatcccatcaggctgttattatgaaggaatttgattg 2412
||||||| ||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
||| ||||||||| ||| ||||||||| ||| ||||||||| ||| ||||||||| ||| ||||||||| |||

Qy 2265 ATTCAATTATTCCTCAGATTGCTGAATCCCATCCCATCAGGCTGTTATTATGAAGGAATTTGATTG 2324

Db 2413 ctttgtgcacagcaggacctgtgcttttgagatatttttccctctcttttaaaatattc 2472
||||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
||| ||||||||| ||| ||||||||| ||| ||||||||| ||| ||||||||| ||| ||||||||| |||

Qy 2325 CTTTGGTCACACAGGACCTGCTGTTTGAGAT-TTTTTTT---TCTCTTTTAAATATC 2380

Db 2473 ttgtaactacaatgatggcaaacg-atgctaaatgactcgattgtacttgaagtaattgc 2531
||||||| ||||||||| ||||| || ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
||| ||||||||| ||| ||||||||| ||| ||||||||| ||| ||||||||| ||| ||||||||| |||

Qy 2381 CTGTAACATAATGATGGTAAGCCATGTAAATGACTGTAAATGACTTGTACTTGGAGTAATTGC 2440

Db 2532 acgtattttcctgtgcataagaagcgcaaggctgttgagaaaaagaattaaaaatttc 2591
|| | |||||| | |||||| ||| | | || || | || || || || || || || || || || || || || ||
|| | |||||| ||| ||| |||||| ||| ||| || || || || || || || || || || || || || || ||

Qy 2441 ACAATTTTTTC-TATGCATAAAAAAATG--ATGCAGCTGTTCAGAA-AACGAAG-TCTTT 2495

Db 2592 ctcatittgtaggaggaatgatggagtttgctctatacctcagtcagtgtgtcatcagaga 2651
||||||| || ||||||||| |||| || || || || || || || || || || || || || || || || || ||
||| || || || || || || || || || || || || || || || || || || || || || || || || ||

Qy 2496 TTCATTTTGCAGAGGAATGATGAAATTTT-TCTGTACTTCAGTATGTGTCAACTGAGA 2554

Db 2652 atcatctgcattagttttaatctttaatgtcaggaatcagattgcaaaagccgagtaatt 2711
|||| | ||||||||| ||||||||| ||||| || |||||| |||||| || || || || || || || || ||
||| || || || || || || || || || || || || || || || || || || || || || || || || ||

Qy 2555 GTCATATACATTAGTTTAAATCTCTTAATATTGAGAATCAGGTTGCCAAACGGATGAGTT 2614

Db 2712 gttatgtgtgggaataataagggaaactaccacactgcctaccaggtctgagaaatagatat 2771
|||| | || || || || || || || || || || || || || || || || || || || || || || || ||
|||| | || || || || || || || || || || || || || || || || || || || || || || || ||

Qy 2615 ATTATCTATGGAAATGTGAGAAATGT-CTA-ATAGCCCCATAAAGTCTGAGAAATAGGTAT 2672

Db 2772 taaagtgccttagggaagctaacagagagagactggaggattgctg 2814
|||| | || || || || || || || || || || || || || || || || || || || || || || || ||
|||| | || || || || || || || || || || || || || || || || || || || || || || || ||

Qy 2673 CAAATAGTTTAGCAAAATGACAGGAGAGAACAGTAGGATTGCTG 2715
```

```
RESULT 4
LOCUS MM1YKMR 2602 bp RNA ROD 06-JUL-1995
DEFINITION M.musculus mRNA for intestinal tyrosine kinase.
ACCESSION 248757
KEYWORDS intestinal tyrosine kinase.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Myomorpha;
Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2602)
AUTHORS Thuveson,M., Albrecht,D., Zurcher,G., Andres,A.C. and Ziemiecki,A.
TITLE iyk, a novel intracellular protein tyrosine kinase differentially
expressed in the mouse mammary gland and intestine
JOURNAL Biochem. Biophys. Res. Commun. 209 (2), 582-589 (1995)
MEDLINE 95251656
REFERENCE 2 (bases 1 to 2602)
AUTHORS Ziemiecki,A.
TITLE Direct Submission
JOURNAL Submitted (21-MAR-1995) to the EMBL/GenBank/DBJ databases.
Ziemiecki A., Laboratory for Clinical and Experimental Research,
Department of Clinical Research, Tiefenaustrasse 120, Berne,
SWITZERLAND, CH-3004
COMMENT NCBI gi: 736263
FEATURES Location/Qualifiers
```

```
source 1..2602
/organism="Mus musculus"
/strain="Balb/c"
/clone="IYK"
/tissue_type="mammary gland"
/cell_type="epithelial"
/cell_line="31D"
/clone_lib="31D-lambda-ZAP"
/sex="Female"
315..1853
/standard_name="IYK"
/note="pid:g736264; NCBI gi: 736264"
/codon_start=1
/function="protein tyrosine kinase"
/citation=[1]
/product="intestinal tyrosine kinase"
/translation="MGSVCVRLWAYLQPFLLPCWSQEADKSVWIENPGAFCPPEAPRSQ
EPERSHGQYFVALFDYOARTAEDLSFRAGDKLQVLDTSHEGWLARHLEKKGTLGQQ
IQGYIPSNYYAEDRSLOAEPWFFCAIKRADAEOILLYSENQTGAFLIRESETQKDFS
LSVLDEGVWKHYRIRRLDEGGFFLTRRKVFSTLINEFVNYVTTTSDGLCVKLEKPCIKI
QVPTPFDLSYKTADQWEIHRNSIQLLKRLGSGQFGEVWEGLMNNTTPVAVKTLKPGSM
DPNDFLREAQIMKSLRHPKLIQLYAVCTLEDPIYIITELMRHGSLOEYLQNDGGSKIH
FIQOVDMAAQVASGMAYLESQNYIHRDLAARNVLVGEHNIYKVADFLARVFKVDNE
IYESKHEIKLPVKWTAPFAIRTNKFSIKSDVWSFGILLIVELIITYGKMPYSGMTGAQVI
QMLSQNYRLPQPSNCPQQFYSIMLECNVPEKQRPFTETLHWKLEDYFETDCSYSDTN
KFIN"
```

```
BASE COUNT 780 a 537 c 596 g 685 t 4 others
ORIGIN

Query Match 41.4%; Score 1148; DB 54; Length 2602;
Best Local Similarity 84.8%; Pred. No. 0.00e+00;
Matches 1536; Conservative 0; Mismatches 253; Indels 23; Gaps 18;
```

```
Db 459 agccatggccagtatatttggctctgtttgattaccaagcacgtactcagagagacctg 518
|| |||||| || |||||||| ||||||||| ||||||||| || || |||| |||||| ||
Qy 489 AGGCATGGCCACTACTTTGTGGCTTTGTTGATTACAGCTCGGACTCGGAGGACTTG 548

Db 519 agcttccgtgccggtgacaaactccaagctcttgacacttcgcattgagggctggtggtg 578
|||||| || ||||||||| ||||| |||||||| |||||||| ||||||||| |||||||
Qy 549 AGCTTCGGAGCAGGTGACAAACTTCAACTCTGGACACTTTGCATGAGGCGTGGTGCTT 608

Db 579 gccagacatttgagagaaggaaggaacggcttaggtcacagctacagggctacatcct 638
|||||| |||||||| | || ||| |||||| ||||| ||||| ||||| |||||
Qy 609 GCCAGACACTTGGAGAAAAAGACGAGATGGCTCCAGTACAGCAACTACAAGGCTATATTCCT 668

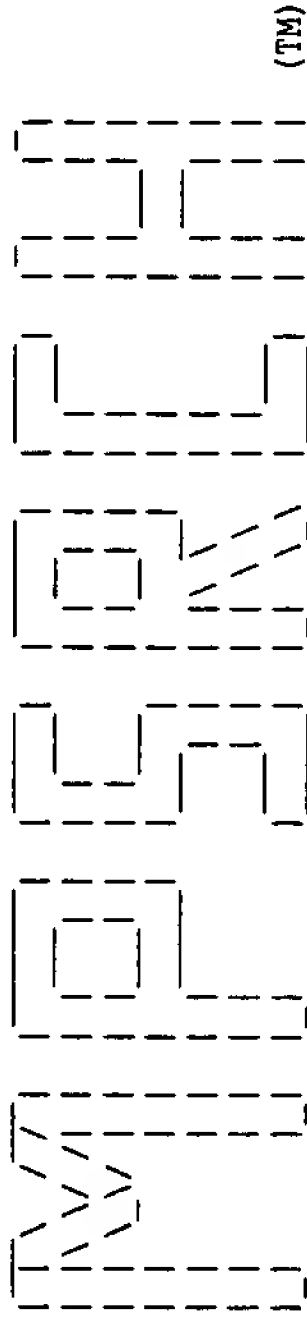
Db 639 tccaattacgtggcgaggaccggaggtctccaggcagagccgtggtttttggagcaatc 698
|| || |||||| ||||| | || || ||||||||| ||||| || |||||||||
Qy 669 TCTAACTACGTGGCTGAGGACAGAAAGCCTACAGCAGAGCCGTGCTTTTGGAGCAATC 728

Db 699 aaaagagcagatgcagaaaaaacaactctgtatttcagaaaaaccagagcggccttcta 758
|||| ||||||||| |||||||| | ||||||||| ||||| || |||||||||
Qy 729 GGAAGATCAGATGCAGAAACAACACTATTATATTCAAAAAACAAGACCCGTTCCTTTCTA 788

Db 759 atcagagagagtgcagaccagaaggtgcacttttccctctcagtttttagatgaagtggtt 818
|||||| ||||| | ||| || || || || || || || || || || || || || || ||
Qy 789 ATCAGAGAAACTGAAAGCCAAAAGGAGAAATCTCTCTTTCAGTTTTCAGATGGAGCAGTT 848

Db 819 gtaaacactacagaataagaaggttggtatgaaggtgcttctcctcaccaggaggaaa 878
||||||| |||||||| | ||| |||||||| || || || |||| | || | ||
Qy 849 GTAAACACTACAGAAATTAAAGACTGCATGAAGGGGATTTTTTCTCAGCGGAAGA 908

Db 879 gtcttttcaaccctgaatgaattcgtgaactactacacacaacaagaagtgcaggctgtgt 938
||||||| ||||| ||||| ||| | |||||| || |||||||| |||||||| ||||||
Qy 909 ATCTTTCAACACTGAACGAATTTGTGAGCCACTACACCAAGACAAGTACGGCCTGTGT 968
```

[illegible]

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on:	Fri May 31 12:05:40 1996;	MasPar time 212.42 Seconds	867.085 Million cell updates/sec
---------	---------------------------	----------------------------	----------------------------------

Tabular output not generated.

```

Title: >US-08-426-509-5
Description: (1-2770) from US08426509.seq
Perfect Score: 2770
N.A. Sequence: 1 CCGACTGGTCGAAGACAG.....AAGTACCAAAAAAAAAAAAA 2770
Comp: GGCCCTCACCAGCTTCTGTC.....TTCATGCTTTTTTTTTTTT

```

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 84802 seqs, 33246950 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq22
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16

Statistics: Mean 10.064; Variance 6.608; scale 1.523

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %		DB	ID	Description	Pred. No.
	Score	Match				
c 1	2619	94.5	7607 16	T03097	Protein tyrosine-kinase	0.00e+00
c 2	2619	94.5	7607 8	Q49754	pTK gene LpTK-2.	0.00e+00
c 3	131	4.7	149 8	Q49749	pTK gene LpTK2 partia	1.76e-60
c 4	131	4.7	149 16	T03092	Protein tyrosine-kinase	1.76e-60
5	102	3.7	1611 7	Q46688	Human pp60 c-src gene	7.57e-43
6	94	3.4	1047 2	Q10572	Human Natriuretic Pep	4.63e-38
7	79	2.9	4550 3	Q14936	Abelson Related Gene,	3.27e-29
8	77	2.8	3437 3	Q14937	Abelson Related Gene,	4.78e-28
9	75	2.7	1254 3	Q13983	Lck gene fused with p	6.92e-27

c	10	73	2.6	1047	2	Q10572	Human Natriuretic Pep	9.92e-26
	11	70	2.5	1602	7	Q46687	Chicken pp60 c-src ge	5.28e-24
	12	53	1.9	2507	13	Q81189	Breast tumour kinase,	1.81e-14
	13	44	1.6	1739	4	Q27539	Tyrosine Kinase recep	1.18e-09
	14	45	1.6	2310	3	Q14851	Clone pTb1283 encodin	3.52e-10
	15	45	1.6	2469	2	Q13311	Basic FGF receptor.	3.52e-10
	16	45	1.6	2662	3	Q21003	flg receptor protein	3.52e-10
	17	45	1.6	2676	3	Q14850	Clone pTb1284 encodin	3.52e-10
	18	45	1.6	2856	2	Q10448	Human basic fibroblas	3.52e-10
	19	45	1.6	3328	3	Q20914	Human bFGF receptor	3.52e-10
	20	45	1.6	3416	3	Q21004	bek receptor protein	3.52e-10
	21	45	1.6	3525	4	Q27658	N-sam cDNA.	3.52e-10
	22	41	1.5	91	9	Q51746	Oligonucleotide probe	4.21e-08
c	23	41	1.5	91	9	Q51746	Oligonucleotide probe	4.21e-08
	24	42	1.5	204	1	N81164	Base substituted E.co	1.29e-08
c	25	42	1.5	4138	10	Q55179	Human orphan receptor	1.29e-08
	26	39	1.4	204	1	N81164	Base substituted E.co	4.39e-07
	27	40	1.4	1590	10	Q56698	Protein-tyrosine-kin	1.36e-07
	28	40	1.4	1601	11	Q62130	Mouse tyrosine kinase	1.36e-07
	29	39	1.4	3429	14	Q85413	Human JAK1 kinase cod	4.39e-07
	30	39	1.4	3540	10	Q26039	JAK1 encoding DNA.	4.39e-07
	31	40	1.4	3760	13	Q75334	Murine tyrosine kinas	1.36e-07
	32	40	1.4	4175	11	Q62129	Mouse tyrosine kinase	1.36e-07
	33	40	1.4	4176	10	Q56697	Protein-tyrosine-kin	1.36e-07
	34	40	1.4	4364	10	Q63730	Protein tyrosine kina	1.36e-07
	35	38	1.4	4544	1	N90355	cDNA encoding platele	1.40e-06
	36	40	1.4	4640	16	Q91999	Mouse tie-2 receptor	1.36e-07
	37	38	1.4	5427	4	Q27447	Type B human platelet	1.40e-06
	38	38	1.4	5719	1	N90388	cDNA encoding human p	1.40e-06
	39	37	1.3	2345	2	Q10868	KGF receptor gene.	4.41e-06
	40	37	1.3	2345	3	Q14049	Human ect gene.	4.41e-06
	41	37	1.3	3069	9	Q52474	Human RYK cDNA clone.	4.41e-06
	42	37	1.3	3754	13	Q84782	Protein-tyrosine-kin	4.41e-06
	43	37	1.3	3962	15	Q92520	Human mammary carcino	4.41e-06
	44	37	1.3	3962	15	Q92522	Human mammary carcin	4.41e-06
	45	35	1.3	5178	3	Q14379	Human insulin recepto	4.26e-05

ALIGNMENTS

RESULT	1	
ID	T03097	standard; DNA; 7607 BP.
AC	T03097;	
DT	14-FEB-1996	(first entry)
DE	Protein tyrosine-kinase LpTK2	gene.
KW	Protein tyrosine-kinase; pTK; LpTK2;	agonist; cell growth;
KW	differentiation; ss.	
OS	Homo sapiens.	
PN	W09527061-A1.	
PD	12-OCT-1995.	
PF	04-APR-1995;	U04228.
PR	04-APR-1994;	US-222616.
PA	(GETH) GENENTECH INC.	
PI	Bennett BD, Goeddel D,	Lee JM, Matthews W, Tsai SP;
PI	Wood WI;	
DR	WPI; 95-366160/47.	
DR	P-PSDB; R85929.	
PPT	Agonist antibodies which activate	specific protein tyrosine
PT	kinase(s) - also activate chimeric	proteins of kinase extracellular
PT	domain and Ig constant domain,	useful for studying, and therapeutic
PT	modulation of, cell growth and	differentiation
PT	Disclosure; Page 48-56; 125pp;	English.
PS	DNA probes based on protein	tyrosine-kinase (pTK) sequences
CC	to screen cDNA libraries to	identify novel pTK genes. A LpTK2
CC	gene	

CC (T03097) was isolated from lymphocytic and megakaryocytic cell
CC libraries. The gene can be used to produce recombinant LpTK2, to
CC identify other new pTK genes, or to design drugs, peptides or
CC antisense constructs that modulate pTK activity.
SQ Sequence 7607 BP; 1954 A; 1851 C; 1693 G; 2109 T;

Query Match 94.5%; Score 2619; DB 16; Length 7607;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;
Matches 2666; Conservative 0; Mismatches 5; Indels 7; Gaps 6;

Db 987 ggtactcttttcttttattactcagaagctagggccacagcaate-tactgttct 1045
|||||
Cp 2757 GGTACTTCTTTTCTTTATTAATTACTCAGAAGCTTAGGCCACAGCAATCCTACTGTTCT 2698
|||||

Db 1046 cctctcaatttctaaactattttgatacctatttctcagactttatgggctattagaca 1105
|||||
Cp 2697 CCTCTCATTTTCCTAACATAATTTTGATACCTATTCTCAGACTTTATGGGCTATTAGACA 2638
|||||

Db 1106 tttctcacatttccatagataataactcatccgttttgcgaacctgattctcaatattaag 1165
|||||
Cp 2637 TTTCACATTTCCATAGATAATAACTCATCCGTTTTTGC AACCTGATTC TCAATATAAG 2578
|||||

Db 1166 agattaaaaactaatgtatagctctcagctgcacacatactgaagtacagaaaaattcca 1225
|||||
Cp 2577 AGATTAAAACTAATGTATATGACTCTCAGCTTGACACATACTGAAGTACAGAAAAATTCCA 2518
|||||

Db 1226 tcatttccctctgcaaatgaaaagacttcgttttctcaacagctgcatcatttttta 1285
|||||
Cp 2517 TCATTTCTCTTCGCAAAATGAAAAGACTTCGTTTTCTCAACAGCTGCATCATTTTTTTA 2458
|||||

Db 1286 tqcatagaaaaaatgtgcaattactccaagtacaatcaagtcatttaacatggctttac 1345
|||||
Cp 2457 TGCATAGAAAAAAATGTGCAATTACTCCAAGTCAAAATCAAGTCATTTAACAATGGCTTTAC 2398
|||||

Db 1346 catcattgtagttacaggatattttaaaagagaaaaaatctcaaaagcacaggtcctg 1405
|||||
Cp 2397 CATCATGTAGTTACAGGATATTTTAAAGACAGAAAAAAATCTCAAGCCACAGGTCCTG 2338
|||||

Db 1406 ctgtgcagcaagcaatcaaatcccttcataataacagcctgatgggattcagcaatctg 1465
|||||
Cp 2337 CTGTGCAGCAAGCAATCAAAATCCTTCATAATAACAGCCTGATGGGATTACGCAATCTG 2278
|||||

Db 1466 aggaataatgaataaaccacttaatcagtaaacaggaaaatgctaca-acag-tcactga 1523
|||||
Cp 2277 AGGAATAATGAATAACCACTCTAAATCAGTAACAGGAAAATGCTACACAGTCATCACTGA 2218
|||||

Db 1524 gtaaaattggactatcatctgttgattctcttgatcg--acatttcaaacaaataaatgg 1581
|||||
Cp 2217 GTAAAAATTGGACTATCATCTGTTGATTCTCTTGATCGCGACATTTCAACAATAAATGG 2158
|||||

Db 1582 aaatgtaagtatctcttaaaagaaaaataaacttggtttagtgtgcttaatttaccagg 1641
|||||
Cp 2157 AAATGTAAGTATCTCTTAAAAACAAAAATAACTTGGTTTAGTGTGCTTAATTTT-CCAGG 2099
|||||

Db 1642 cagtgaggaaattatatatcaccttgactgtcctgcagtggtggccagtcacaataaatgc 1701
||||| || |||||
Cp 2098 CAGTGAGCAA-TTATATAGACCTTGACTGTCTCTGCAGTGTGGCCAGTCAAATAAATGC 2040
|||||

Db 1702 acaataatctttttcataatacatatggccaaactttatcctatacacttggaatgtcagga 1761
|||||
Cp 2039 ACAATAATCTTTTTTCATAATACATGGCCAACTTTATCCTATCACTTGAATATGTCAGGA 1980
|||||

Db 1762 taaactgattgtgcagttggttgataaacattgtattttggaatggattatttgaatttgt 1821
|||||

Cp 1979 TAAACTGATTGTGCAGCTGGFTTGATAACATTGTAATTTTGGAAATGGATTATTTGAAATTTGT 1920
|||||

Db 1822 ttgtctactttattatttgatattcttctccagtggtcatcttatgaagttatttgcac 1881
|||||
Cp 1919 TTTGCTACTTTATTTGATATTTCTTCTCCAGTGTCTAICTTTATGAAGTTATTTGCAATC 1860
|||||

Db 1882 tgaatatgaagagctgttttcaaaaatagcttccaagtttccaacgcagtgctctcaaatgt 1941
|||||
Cp 1859 TGAATATGAAGAGCTCTGTTTCAAAAATAGTCTTCAAGTTTCCAACGGCAGTGTCTCAAAATGT 1800
|||||

Db 1942 aggtcgttccttaggctctgcattccagcactccaacatgatgttgtaaaattgctgtgg 2001
|||||
Cp 1799 AGGTGGTTCCTTAGGCCTGCGATTCCAGCACTCCAACATGATGTTGTAAAAATTTGCTGTGG 1740
|||||

Db 2002 acagttgggatggttgcggaagctctatagtttgcgccaacatctggattaccctgggcacc 2061
|||||
Cp 1739 ACAGTTGGATGGTTGGGAAAGTCTATAGTTTTCAGCCCAACATCTGGATTACCTGGGCACC 1680
|||||

Db 2062 tgtcataccactgtaaggcatttttgcataaagtaaatgatttcataaaagaaggattccaaa 2121
|||||
Cp 1679 TGTCTATACCACCTGTAGGCAATTTTGCCATAAGTAATGATTTTCATAAAGAAGGATTCCTAAA 1620
|||||

Db 2122 tgaccatacatcggaacttaatgctgaatttattactacgaatggcttcggcgcgagtc 2181
|||||
Cp 1619 TGACCATACATCGGACTTAATGCTGAATTTATTACTACGAATGGCTTCGGGGCGCATCTCA 1560
|||||

Db 2182 cttcacoggcagcctttatttcgtgctcagattcctcattatctcattatctacctttaa 2241
|||||
Cp 1559 CTTTACCGGGCAGCTTTATTTTCGTGTCAGATTCTAGATGCTCTCATTTATCTACCTTTAAA 1500
|||||

Db 2242 aactctgggcaagtccaaaatctgctacttttgcagatattatggtcaccaacgaggacatt 2301
|||||
Cp 1499 AACTCTGGCAAGTCCAAAAATCTGCTACTTTTGTAGATATTATGTTTACCACACGAGGACATT 1440
|||||

Db 2302 tctggcagccagatctctgtgaatgtagttccgagagactccagataggccattccagaggc 2361
|||||
Cp 1439 TCTGGCAGCCAGATCTCTGTGAATGTAGTTCCGAGACTCCAGATAGGCCAATCCAGAGGC 1380
|||||

Db 2362 aacctgtgccccatgtctacctgttgagtcagatggatttttgcagtgctcattttg 2421
|||||
Cp 1379 AACCTGTCCCGCCATGTCTACCTGT TGAGTCAAGATGGATTTTTCATCCAGTGTCA TTTTG 1320
|||||

Db 2422 gagatattcttgcagacttccatgtctcatcaactctgtaataataataaattggatcttc 2481
|||||
Cp 1319 GAGATATTCTTGCAGACTTCCATGCTCATCAACTCTGTAAATAATATAAATTTGGA TCTTC 1260
|||||

Db 2482 taaagtgcacaacagcataaaagctggataagctttggatgtcttaggttcttcattatctg 2541
|||||
Cp 1259 TAAAGTGCAAAACAGCATAAAGCTGGATAAGCTTTGGATGCTCTTAGGTTCTTCATTATCTG 1200
|||||

Db 2542 tgcctccctcagggaagtcattttggatccattgaacctgggttttaatgttttcactgctac 2601
|||||
Cp 1199 TGCCCTCCCTCAGGAAGTCATTTGGATCCATTTGAACCTGGTTTTTAATGTTTTTCACTGCTAC 1140
|||||

Db 2602 tggagtggtattgttccacagaccttcccatacttcgccaacctgaccagatcccaatcg 2661
|||||
Cp 1139 TGGAGTGGTATTGTTCCACAGACCTTCCCATACTTCGCCAAAAC TGACCAGATCCCAATCG 1080
|||||

Db 2662 cttcagaagctgtatggagttgcggtctctatctcccatgggtccacgggttttatacgaca 2721
|||||
Cp 1079 CTTTCAGAAAGCTGATGGAGTTGGGGTCTATCTCCCATTTGGTCCAGGTTTTTATACGACAA 1020
|||||

Db 2722 atcaaatggagctgggacctggatctttaagcatggtttccccagcttgacacacagggcc 2781
|||||

```
Cp 1019 ATCAATGGAGCTGGGAGCTGGATCTTTAAGCATGGTTTCCCCAGCTTGACACACAGGCC 960

Db 2782 gtcacttgcttggtgtagtggtcacaaattcgttcagttggtgaaagattcttcttcg 2841
|||||
Cp 959 GTCAC TTGCTCTGGGTAGTGGCTCACAAATTCGTTTCAGTGTGGAAAGATCTTCTCTCG 900
GTCAC TTGCTCTGGGTAGTGGCTCACAAATTCGTTTCAGTGTGGAAAGATCTTCTCTCG 900

Db 2842 cgtgagaaaaatcccccttcaccagtcctttaaattcgtagtggttttaacaactgctcc 2901
|||||
Cp 899 CGTGAGAAAAATCCCCCTTCATCCAGTCTTTTAAATTCGTAGCTTTTACAACTGCTCC 840
CGTGAGAAAAATCCCCCTTCATCCAGTCTTTTAAATTCGTAGCTTTTACAACTGCTCC 840

Db 2902 atctaaaactgaaagagaaattctccttttggctttcactttctctgattagaagga 2961
|||||
Cp 839 ATCTAAACTGAAAGAGAGAAATTCCTCTTTTGGCTTTTCAC TTCTCTGATTAGAAGGA 780
ATCTAAACTGAAAGAGAGAAATTCCTCTTTTGGCTTTTCAC TTCTCTGATTAGAAGGA 780

Db 2962 accggtctgtttctgaataataatagttgtttctctgcatctgatcttcogattgctcc 3021
|||||
Cp 779 ACCGGTCTGTTTCTGAATATAATAGTTGTTTCTCGCATCTGATCTTCGGATTGCTCC 720
ACCGGTCTGTTTCTGAATATAATAGTTGTTTCTCGCATCTGATCTTCGGATTGCTCC 720

Db 3022 aaagaaccacgctctgcctgtaggctctgtcctcagccacgtagttagaaggaatata 3081
|||||
Cp 719 AAAGAACCCAGGCTCTGCCGTGTAGGCTTCTGCTCAGCCACGCTAGTTAGAAGGAATA 660
AAAGAACCCAGGCTCTGCCGTGTAGGCTTCTGCTCAGCCACGCTAGTTAGAAGGAATA 660

Db 3082 gcccttagtgctgactggagccatctcgtcttctccaagtgtctggcaaacaccca 3141
|||||
Cp 659 GCCTGTAGTTGCTGACTGGAGCCATCTCGCTCTTTCTCCAAGTGTCTGGCAACACCCA 600
GCCTGTAGTTGCTGACTGGAGCCATCTCGCTCTTTCTCCAAGTGTCTGGCAACACCCA 600

Db 3142 gccctcatgaaagtccagaacttgaagttgtcacctgctcggaagctcaagtcctc 3201
|||||
Cp 599 GCCCTCATGCAAGTGTCCAGAACTTGAAGTTGTCACTGCTCGGAAGCTCAAGTCTC 540
GCCCTCATGCAAGTGTCCAGAACTTGAAGTTGTCACTGCTCGGAAGCTCAAGTCTC 540

Db 3202 agcagtcagcgtggttaatacaaaagccacaagttagtggccatgctctgtgactg 3261
|||||
Cp 539 AGCAGTCCGAGCCTGGTAATCAAAACAGCCACAAGTAGTGGCCATGCCCTCTGTGACTG 480
AGCAGTCCGAGCCTGGTAATCAAAACAGCCACAAGTAGTGGCCATGCCCTCTGTGACTG 480

Db 3262 gggagagcaaggccctggattttcaatcacgggtgactgtctgcctccgtggacaa 3321
|||||
Cp 479 GGGAGAGCAAAAGGCCCTGGATTTTCAATCACGGTTGACTTGTCTGCCCTCCGTGGACAA 420
GGGAGAGCAAAAGGCCCTGGATTTTCAATCACGGTTGACTTGTCTGCCCTCCGTGGACAA 420

Db 3322 acaggggagatagggttctaggtactcccagagcctctgacagatggtgctcattgtgcc 3381
|||||
Cp 419 ACAGGGGAGATAGGGTCTTAGGTACTCCCAGAGCCTCTGCACAGATGTTGCTCATTTGTGCC 360
ACAGGGGAGATAGGGTCTTAGGTACTCCCAGAGCCTCTGCACAGATGTTGCTCATTTGTGCC 360

Db 3382 ttggtgggagaagagcagggttctccctctcccttagtctctgcatccacctt 3441
|||||
Cp 359 TTGGTGGGAGAAGAGAGCAGCGGCTTCTCCCTCTCCCTTAGTCTCTGCGATCCACCTT 300
TTGGTGGGAGAAGAGAGCAGCGGCTTCTCCCTCTCCCTTAGTCTCTGCGATCCACCTT 300

Db 3442 atcttcttcaccaggcaactttgaagtcagcaccacactcaccatacttcggagagtatg 3501
|||||
Cp 299 ATCTTCCTTCACCAGCGCACTTTGAAGTCAGCACCAACTCACCA TACTTCGGAGAGTATG 240
ATCTTCCTTCACCAGCGCACTTTGAAGTCAGCACCAACTCACCA TACTTCGGAGAGTATG 240

Db 3502 caaagtccttcagatcagtcagcagctgggttgagcaaacctcaccatacttcggagagac 3561
|||||
Cp 239 CAAAGTCCCGTTTCAGATCAGTCCAGCAGCTGGGTTGCGAGCAAGTCTTACCTGGAGAGAC 180
CAAAGTCCCGTTTCAGATCAGTCCAGCAGCTGGGTTGCGAGCAAGTCTTACCTGGAGAGAC 180

Db 3562 ttaccggcttgcttctgtggctggaggtgtaccceggaggaactgagcaggagctg 3621
|||||
Cp 179 TTACCGGCTTGCCTTCTGTGCTGGAGGTGTACCCCGAGGCCAAA ACTGAGCAGGAGCTG 120
TTACCGGCTTGCCTTCTGTGCTGGAGGTGTACCCCGAGGCCAAA ACTGAGCAGGAGCTG 120

Db 3622 ggcagctgtcactaggaaggtgtcttttcttcttacc 3659
|||||
Cp 119 GGCAGCTGCTCACTAGGAAGGTGCTTTTCTTCTTATC 82
GGCAGCTGCTCACTAGGAAGGTGCTTTTCTTCTTATC 82
```

```
ID Q49754 standard; DNA; 7607 BP.
AC Q49754;
DT 10-MAR-1994 (first entry)
DE PTK gene lptk-2.
KW PTK; protein tyrosine kinase; catalytic domain; c-kit; megakaryocyte;
KW lymphocyte; amplification; primer; polymerase chain reaction; PCR; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1858..3375
FT /*tag= a
PN W09315201-A.
PD 05-AUG-1993.
PF 22-JAN-1993; U00586.
PR 22-JAN-1992; US-826935.
PA (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
PI Avraham H, Cowley S, Groopman J, Scadden D;
DR WPI; 93-320330/40.
DR P-PSDB; R41941.
PT New protein tyrosine kinase genes and proteins encoded by genes --
PT are of human mega-karyocytic origin
PS Claim 2; Fig 5; 60pp; English.
CC PTK genes were identified using two sets of degenerative
CC oligonucleotide primers: a first set which amplifies all pTK DNA
CC segments (Q49743-44), and a second set which amplifies highly
CC conserved sequences present in the catalytic domain of the c-kit
CC subgroup of pTKs (Q49745-46). The pTK genes identified are described
CC in Q49747-57 and R41897-02.
CC The lptks are expressed in lymphocytic cells, as well as
CC megakaryocytic cells. The partial and full-length lptk2 gene
CC sequences are given in Q49749 and Q49754 respectively. The
CC protein sequence corresp. to Q49749 is claimed (claim 7) and
CC stated as given in the specification, however is missing from
CC the publication.
SQ Sequence 7607 BP; 1953 A; 1851 C; 1694 G; 2109 T;

Query Match 94.5%; Score 2619; DB 8; Length 7607;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;
Matches 2666; Conservative 0; Mismatches 5; Indels 7; Gaps 6;

Db 987 ggtactctcttttcttattactacagaagtctaggccacagcaatc-tactgttct 1045
|||||
Cp 2757 GGTACTTCTTTTCTTTATTATTACTCAGAAAGCTAGGCCACACCAATCCTACTGTTCT 2698
GGTACTTCTTTTCTTTATTATTACTCAGAAAGCTAGGCCACACCAATCCTACTGTTCT 2698

Db 1046 cctctcaatttcctaaactatttgatacctatttctcagacatttatgggtattagaca 1105
|||||
Cp 2697 CCTCTCATTTTCTTAAC TATTTTGATACCTATTTCTCAGACTTATGGGCTATTAGACA 2638
CCTCTCATTTTCTTAAC TATTTTGATACCTATTTCTCAGACTTATGGGCTATTAGACA 2638

Db 1106 ttctcacatttcctcatagataataaactcatccgttttgcacacctgattctcaatattaag 1165
|||||
Cp 2637 TTTCTCACATTTCCATAGATAATACTCATCCGTTTTCGAACCTGATTCTCAATATTAG 2578
TTTCTCACATTTCCATAGATAATACTCATCCGTTTTCGAACCTGATTCTCAATATTAG 2578

Db 1166 agattaaaaactaatgtatgactctcagttgacacatactgaagtacagaaaaattcca 1225
|||||
Cp 2577 AGATTAAAAC TAAATGTATATGACTCTCAGTTGCACATACTGAAGTACAGAAAAATTCCA 2518
AGATTAAAAC TAAATGTATATGACTCTCAGTTGCACATACTGAAGTACAGAAAAATTCCA 2518

Db 1226 tcatttcctctgcaaaatgaaaaagacttcgttttctcaacagctgcatcatttttta 1285
|||||
Cp 2517 TCATTTCC TCTGCAAAATGAAAAAGACTTCGTTTCTCAACAGCTGCATCATTTTTTA 2458
TCATTTCC TCTGCAAAATGAAAAAGACTTCGTTTCTCAACAGCTGCATCATTTTTTA 2458

Db 1286 tgcataaaaaaaaaatgtgcaattactccaagtacaaatcaagtcatttaacatggctttac 1345
|||||
Cp 2457 TGCATAGAAAAAAATCTGCAATTACTCCAAGTACAATCAAGTCA TTTTAACATGGCTTAC 2398
TGCATAGAAAAAAATCTGCAATTACTCCAAGTACAATCAAGTCA TTTTAACATGGCTTAC 2398

Db 1346 catcattgtagttacaggatatttttaaagagaaaaaaatctcaaacacaggtcctg 1405
```

Cp 2397 CATCATGTGTACTTACAGGATATTTTAAAGAGAAAAAATAATCTCAAAGCACAGGTCCTG 2338
Db 1406 ctgtgcagcaagcaatcaaaattccbtcataataacagcctgatgggattcagcaaatctg 1465
Cp 2337 CTGTGCAGCAAGCAATCAAATTCCTTCATATAAFAACAGCCTGATGGATTGAGCAATCTG 2278
Db 1466 aggaataatgaataaaccactctaattcagtaaacagggaaaatgctaca-acag-tcaactga 1523
Cp 2277 AGGAATAATGAATAACCACTCTAATCAGTAACAGGAAAAATGCTACACAGTCATCCTACTGA 2218
Db 1524 gtaaaaattggactatcatctgttgattctcttgatcg--acatttcaaaaataaatgg 1581
Cp 2217 GTAAAAATTGGACTATCATCTCTTGATTCTCTTCATCGGCACATTTCAAACAATAAATGG 2158
Db 1582 aaatgtaagtatctcttaaaaagaaaaaacttggtttagtgccttaattttaccagg 1641
Cp 2157 AAATCTAAGTATCTCTTAAAAAGAAAAATAACTTGGTTAGTGTGCTTAATTTT-CCAGG 2099
Db 1642 cagtgaggaaattatatacatcaccttgactgtcctgcagtggtggttgcacagtcaataaaatgc 1701
Cp 2098 CAGTGAGCAA-TTATAATAGACCTTGACTGTCTCGCAGTGTGGCCAGTCAATAAATGCG 2040
Db 1702 acaataaatcttttccataataacatgagcaactttatccatcatcacttgatataatgtcagga 1761
Cp 2039 ACAAATAATCTTTTTCATAATAATACATGGCCAACTTTATCCTATCACTTGAATATGTCAGGA 1980
Db 1762 taaactgattgtgcagttggttgataaacattgtattttggaatggattatttgaatttgt 1821
Cp 1979 TAAACTGATTGTGCAGTGGTTGATGAATAACATTGTATTTTGAAATGGATTAATTGAATTTGT 1920
Db 1822 tttgctactttattatttgatattcttctccagtggtccatctttatgaagttatttgcac 1881
Cp 1919 TTTGCTACTTTATTAATTGATATTCTTCCAGTGTTCATCTTATGAAGTATATTGCAATC 1860
Db 1882 tgaatatgaagagctgtttccaaaatagctcttcaagtttccaaagcagtgctctcaaatgt 1941
Cp 1859 TGAATATGAAGAGTCTCTTTCAAATACTCTTCAAGTTTCCAAGCAGTCTCTCAAATGT 1800
Db 1942 aggtcgcttcccttaggtctgcattccagcacctccaacatgatgttgtaaaattgctgtgg 2001
Cp 1799 AGGTCTCTCCTTAGGCTCTGCATTCCAGCACTCCAACATGATGTTGTAAAATTGCTGTGG 1740
Db 2002 acagttggatggttgcggaagtctatagttttgagcccaacatctggattaccctgggcacc 2061
Cp 1739 ACAGTGGATGGTTGCGGAAGTCTATAGTTTTCAGCCCAACATCTGGATTACCTGGGCACC 1680
Db 2062 tgtcataccactgtaaaggcattttgcccataaagtaatgatttccataaagaaggattccaaa 2121
Cp 1679 TGTCTACCACTGTAAGGCATTTTGGCATAAGTAATGATTTTCATAAAGAGGATTCCAAA 1620
Db 2122 tgaccatacatcggacttaatgctgaatttattactacgaatggcttcgggcgcagtccea 2181
Cp 1619 TGACCATACATCGGACTTAATGCTGAATTTATTACTACGAATGGCTTCGGGCGCAGTCCA 1560
Db 2182 cttcacccggcagctttatttcgtgctagattcatagatgtcttcattatctaccttaaa 2241
Cp 1559 CTTACCGGCGAGCTTTATTTCCGTGCTAGATTTCATAGATGTCTTCATTATCTACCTTAAA 1500
Db 2242 aactctggcaagtccaaaactgctactttgtagatattatgttcaccaacgaggacatt 2301
Cp 1499 AACTCTGGCAAGTCCAAAACTGCTACTTTGTAGATATATATTGTTCAACCAACGAGGACATT 1440
Db 2302 tctggcagccagatctctgtgaatgtagttccgagactccagataggccattccagaggc 2361

Cp 1439 TCTGGCAGCCACGATCTCTGTGAATGTAGTTCGGAGACTCCAGATAGGCCATTCCAGAGGC 1380
Db 2362 aacctgtccgcctcatgtctacctgttgagtcagatggatttttgatccagtgctcattttg 2421
Cp 1379 AACCTGTCCGCGCATGCTACCTGTTGAGTTCAGATGGATTTTGTGATCCAGTGTCTATTG 1320
Db 2422 gagataattcttcgagacttccatgctcatcaactctgtaaatctgtaataataaatggatcttc 2481
Cp 1319 GAGATATTCTTTCAGACTTCCATGCTCTCATCAACTCTGTAAATAATATAAATTGGAATCTC 1260
Db 2482 taaagtgcacaacagcataaaagctggataagcttttgatgctcttaggttcttcattatctg 2541
Cp 1259 TAAAGTGCAAACAGCATAAAGCTGGATFAAGCTTTGGATGCTCTTAGGTTCTTCATTATCTG 1200
Db 2542 tgctccctcaggaagtcatttggatccattgaacctggttttaaattgttttcactgctac 2601
Cp 1199 TGCCTCCCTCAGGAAGTCATTTGGATCCATTGAACCTGGTTTAAATGTTTTCACTGCTAC 1140
Db 2602 tggagtggtaattgttccacagaccttccatacttgcgaactgcacagatccccaatcg 2661
Cp -1139 TGGAGTGGTATTGTTCCACAGACCTTCCCATACTTCGCCAAACTGACCAGATCCCATACTG 1080
Db 2662 cttcagaagctgtatggagttgcggtctatctccattggctccacggttttatacagaaa 2721
Cp 1079 CTTCAAGAGCTGATGGAGTTGCGGTCTATCTCCCATTTGGTCCACGGTTTATACGACAA 1020
Db 2722 atcaaatggagctgggacctggatctttaagcatggtttccccagcttgcacacacagggcc 2781
Cp 1019 ATCAAATGGAGCTGGGACCTGGATCTTTTAAAGCATGGTTTCCCGAGCTTGACACACAGGCC 960
Db 2782 gtcacttgtcttggtagtggtgcacaaaattcgttcagtggtgaagaagattcttcttcg 2841
Cp 959 GTCACTTGCTTGGTGTAGTGGCTCACAAAATTCGTTCACTGTTGAAAAGATTCTTCTTCG 900
Db 2842 cgtgagaaaaatcccccttccatccagtccttttaattctgtagtggtttacaactgctcc 2901
Cp 899 CGTGAGAAAAATCCCCCTTCATCCAGTCTTTTAAATTTCTGTAGTGTTTTACAACATGCTCC 840
Db 2902 atctaaaaactgaagagagaaattctccttttggcttccacttctctgattagaaagga 2961
Cp 839 ATCTAAAACTGAAAGACAGAAATTCCTCTTTTGGCTTTTCACTTCTCTGTAGTAAAGGA 780
Db 2962 accggtcttgttttctgaataataatagttggttctctgcatctgatcttcogattgctcc 3021
Cp 779 ACCGGTCTTGTTTTCTGAAATATAATAGTTGTTTCTCTGCACTCTCGGATTGCTCC 720
Db 3022 aaagaaccacggctctgcctgtaggcttctgctcagccacgtagttaagaaggaaatata 3081
Cp 719 AAAGAACCAACGGCTCTGCCCTGTAGGCTTCTGCTCAGCCACCGTAGTTAGAAGGAATATA 660
Db 3082 gcctttagttgctgactggagccatctcgcttttctccaagtgtctggcaaacaccaca 3141
Cp 659 GCCTTGTAGTTGCTGACTGGAGCCA TCTCGTCTTTTCTCCAAGTCTCTGCAAAACCCACCA 600
Db 3142 gccctcatgeaaaagtdcccagaacttgaagtttgtcacctgctcgggaagctcaagtcctc 3201
Cp 599 GCCCTCATGCAAAAGTGTCCAGAACTTGAAGTTTGTCACTGCTCGGAAGCTCAAGTCCCTC 540
Db 3202 agcagtcgagccttggttaatacaaaagccacaagtagtggccatgcctctgtgactg 3261
Cp 539 AGCACTCCGAGCCCTGGTAAATCAAAACAAAGCCAAAGTAGTGGCCATGCCCTCTGTGACTG 480
Db 3262 gggagagcaagggccccctggattttcaatcacgggttgacttgtctgcctccgtggacaa 3321

May 31 12:00

US-08-426-509-5.mg

9

Cp 479 GGGAGACAAAGGGCCCTGGATTTTCAATCACGGTTGACTTGCTGCCCTCGTGCAAA 420
Db 3322 acaggggagataggttctaggtactccagagcctctgacagatggtgtcattgtgcc 3381
Cp 419 ACAGGGAGATAGGGTTCTAGGTACTCCAGAGCCTCTGACAGATGTTGCTCAFTTGCC 360
Db 3382 ttggtgggagaagagagcaggccttccctccctctcccttagtctctgcgatccacctt 3441
Cp 359 TTGCTGGGAGAGAAGAGCAGGGCTTCTCCCTCTCCCTTAGTCTCTGCGATCCACCTT 300
Db 3442 atcttcttcaccaggcaactttgaagtcagcaccacactcaccatacttcggagagtatg 3501
Cp 299 ATCTTCTTCACGAGCAACTTTGAAGTCAGCAACCAACTCACCATACTTCGGAGACTATG 240
Db 3502 caaagtcocgttcagatcagtcagcagcagctgggtgcaagtcactcactcgagagac 3561
Cp 239 CAAAGTCCCGTTTCAGATCAGTCAGCAGCAGCTGGGTGCGACCAAGTCCCTACTGGAGAGAC 180
Db 3562 ttaccggcttgctttctgtgctggagtgctaccocggaggcaaaactgagcaggagctg 3621
Cp 179 TTACCGGCTTGCTTTCTGTGGCTGGAGGTGCTACCCCGAGGCAAAACTGAGCAGGAGCTG 120
Db 3622 ggcagctgcactaggaagtgcttttcttcttctatc 3659
Cp 119 GGCAGCTGCTCACTAGCAAGGTGCTTTTCTTCTTATC 82

RESULT 3

ID Q49749 standard; DNA; 149 BP.
AC Q49749;
DT 10-MAR-1994 (first entry)
DE pTK gene LpTK2 partial sequence.
KW pTK; protein tyrosine kinase; catalytic domain; c-kit; megakaryocyte;
KW lymphocyte; amplification; primer; polymerase chain reaction; PCR; ss.
OS Homo sapiens.
PN W09315201-A.
PD 05-AUG-1993.
PF 22-JAN-1993; U00586.
PR 22-JAN-1992; US-826935.
PA (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
PI Avraham H, Cowley S, Groopman J, Scadden D;
DR WPI; 93-320330/40.
PT New protein tyrosine kinase genes and proteins encoded by genes -
PT are of human mega-karyocytic origin
PS Claim 2; Fig 3A; 60pp; English.
CC pTK genes were identified using two sets of degenerative
CC oligonucleotide primers: a first set which amplifies all pTK DNA
CC segments (Q49743-44), and a second set which amplifies highly
CC conserved sequences present in the catalytic domain of the c-kit
CC subgroup of pTKs (Q49745-46). The pTK genes identified are described
CC in Q49747-57 and R41897-02.
CC The LpTKs are expressed in lymphocytic cells, as well as
CC megakaryocytic cells. The partial and full-length LpTK2 gene
CC sequences are given in Q49749 and Q49754 respectively. The
CC protein sequence corresp. to Q49749 is claimed (claim 7) and
CC stated as given in the specification, however is missing from
CC the publication.
SQ Sequence 149 BP; 37 A; 37 C; 24 G; 51 T;
Query Match 4.7%; Score 131; DB 8; Length 149;
Best Local Similarity 95.2%; Pred. No. 1.76e-60;
Matches 138; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

May 31 12:00

US-08-426-509-5.mg

10

Db 5 gaattccttcggcgccatccatttcaccggcagcgtttatttcgtctagattcataga 64
Cp 1581 GAATGGCTTCGGCGCAGTCCACTTCACGGCAGCTTTATTTCGTCTAGATTTCATAGA 1522
Db 65 tgtcttcattatctaccttaaaactctggcaagtcctcaaaatctgctactttgtagatat 124
Cp 1521 TGTCTTCATTATCTACCTTAAAAACTCTGGCAAGTCCAAAATCTGCTACTTTGTAGATAT 1462
Db 125 tatgttcaccaacgaggacattcct 149
Cp 1461 TATGTTCAACAACGAGGACATTCT 1437

RESULT 4

ID T03092 standard; DNA; 149 BP.
AC T03092;
DT 14-FEB-1996 (first entry)
DE Protein tyrosine-kinase LpTK2 DNA fragment.
KW Protein tyrosine-kinase; pTK; LpTK2; agonist; cell growth;
KW differentiation; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..60
FT /*tag= a
PN W09527061-A1.
PD 12-OCT-1995.
PF 04-APR-1995; U04228.
PR 04-APR-1994; US-222616.
PA (GETH) GENENTECH INC.
PI Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;
PI Wood WI;
DR WPI; 95-366160/47.
DR P-PSDB; R85931.
PT Agonist antibodies which activate specific protein tyrosine
PT kinase(s) - also activate chimeric proteins of kinase extracellular
PT domain and Iq constant domain, useful for studying, and therapeutic
PT modulation of, cell growth and differentiation
PS Disclosure; Page 36; 125pp; English.
CC DNA probes based on protein tyrosine-kinase (pTK) sequences were used
CC to screen cDNA libraries to identify novel pTK genes. A LpTK2 gene
CC fragment (T03092) was isolated from lymphocytic and megakaryocytic
CC cell line libraries and encoded a peptide (R85931) showing homology
CC to known pTKs. The gene fragment can be used to identify other
CC new pTK genes, or to design drugs, peptides or antisense constructs
CC that modulate pTK activity.
SQ Sequence 149 BP; 37 A; 37 C; 24 G; 51 T;

Query Match 4.7%; Score 131; DB 16; Length 149;
Best Local Similarity 95.2%; Pred. No. 1.76e-60;
Matches 138; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 5 gaattccttcggcgccatccatttcaccggcagcgtttatttcgtctagattcataga 64
Cp 1581 GAATGGCTTCGGCGCAGTCCACTTCACCGGCAGCTTTATTTCGTCTAGATTTCATAGA 1522
Db 65 tgtcttcattatctaccttaaaactctggcaagtcctcaaaatctgctactttgtagatat 124
Cp 1521 TGTCTTCATTATCTACCTTAAAAACTCTGGCAAGTCCAAAATCTGCTACTTTGTAGATAT 1462
Db 125 tatgttcaccaacgaggacattcct 149
Cp 1461 TATGTTCAACAACGAGGACATTCT 1437

May 31 12:00

US-08-426-509-5.mg

11

RESULT 5

ID Q46688 standard; cDNA to mRNA; 1611 BP.
AC Q46688;
DT 23-DEC-1993 (first entry)
DE Human pp60 c-src gene.
KW Endothelial; tyrosine kinase protein; pp60 c-src; ss.
OS Homo sapien.
PN W09314193--A.
PD 22-JUL-1993.
PF 05-JAN-1993; U500445.
PR 06-JAN-1992; U5-820011.
PA (UYVA) UNIV YALE.
PI Bell L, Luthringer DJ, Madri JA, Warren SL;
DR WPI; 93-243209/30.
DR P-PSDB; R39705.

PT Genetically engineered endothelial cells - which exhibit enhanced
PT cell migration, urokinase-type plasminogen activator activity,
PT and reduced mononuclear cell adhesion and fibronectin prodn
PS Disclosure; Page 69-72; 91pp; English.
CC The DNA encoding a portion or (more preferably) the entire pp60
CC c-src polypeptide is used to transform endothelial cells.
CC Transformed cells produce increased amounts of pp60 c-src and have
CC improved therapeutic properties. They migrate at faster rates than
CC non-transformed counterparts; have an enhanced ability to inhibit
CC the formation of thrombi and/or dissolve thrombi once they have
CC formed and exhibit reduced mononuclear cell adhesion. They can also
CC be used to improve the success of surgical procedures such as
CC coronary angioplasty, heart bypass surgery, vessel graft and stent
CC implantation.

SQ Sequence 1611 BP; 334 A; 507 C; 504 G; 266 T;

Query Match 3.7%; Score 102; DB 7; Length 1611;
Best Local Similarity 64.8%; Pred. No. 7.57e-43;
Matches 256; Conservative 0; Mismatches 136; Indels 3; Gaps 2;

Db 841 tttagcgaggtgtgatggggacctggaacgtgacgtaccacaggtggccatcaaaacccctg 900
||||||| || ||| || ||||| ||||| || || ||||| |
QY 1098 TTTGGGAAGTATGGGAAGCTCTGTGGAACAATACCATTCCAGTAGCACTGAACATTA 1157

Db 901 aagcctggcacgatgtctccagaggccttcctgcaggagcccgatgaagaagctg 960
|| || || || || || || || || || || || || || || || || || || || || || ||
QY 1158 AAACAGGTTCAATCGATCCAAATGACTTCCTCGAGGCGCACAGATATGAAGACCTA 1217

Db 961 aggcagtgaagctggtgcagttgtatgtgtg-gt--ttcacaggagcccattacatc 1017
|| ||| ||||| || || || || || || || || || || || || || || || || || || ||
QY 1218 AGACATCCAAGCTTATCCAGCTTTATGCTGTTGCACCTTTAGAAGATCCAATTATATT 1277

Db 1018 gtcacggagtacatgagcaaggggagttgtgactttctcaagggggagacaggcaag 1077
| || |||| ||||| || || || || || || || || || || || || || || || || || ||
QY 1278 ATTACAGAGTTGATGAGACATGGAAGCTCTGCAAGAAATATCTCCAAATGACACTGGATCA 1337

Db 1078 tacctgcggtgcctcagctggtggacatggtgctcagatgcctcaggcattgcgtac 1137
| | | ||| ||||| || || ||||| || || || || || || || || || || || || || ||
QY 1338 AAAATCCATCTGACTCAACAGGTAGACATGGGGGCACAGGTTGCCCTCTGGAATGGCCTAT 1397

Db 1138 gtggagcggatgaactacgtccacgggaccttcgtgcagccaacatcctggtgggagag 1197
||||| ||||| || || || || || || || || || || || || || || || || || || ||
QY 1398 CTGGAGTCTCGGAAGTACATTACACAGATCTGGCTGCCAGAAATGCTCCTCGTTGGTGAA 1457

Db 1198 aacctggtgtgcaaagtgccgacttttggctggc 1232
| | | ||||| || || ||||| || ||
QY 1458 CATATAATCTACAAGTAGCAGATTTTGGACTTGC 1492

May 31 12:00

US-08-426-509-5.mg

12

RESULT 6

ID Q10572 standard; DNA; 1047 BP.
AC Q10572;
DT 09-APR-1991 (first entry)
DE Human Natriuretic Peptide Receptor B.
KW NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW hyperaldosteronism; glaucoma; guanyl cyclase.
OS Homo sapiens.

FH Key Location/Qualifiers
FT Peptide 1..22
FT /label= signal sequence
FT Protein 12
FT /label= mature NPBR
FT Domain 23..455
FT /label= extracellular domain
FT /note= "binds natriuretic peptides A,B and C]"
FT Domain 456..456
FT /label= transmembrane domain
FT Domain 479..1047
FT /label= cytoplasmic domain
FT /note= "GC and protien kinase activity"
FT Modified -site 24..26
FT /label= N-glycos site
FT Modified -site 35..37
FT /label= N-glycos site
FT Modified -site 161..163
FT /label= N-glycos site
FT Modified -site 195..197
FT /label= N-glycos site
FT Modified -site 244..246
FT /label= N-glycos site
FT Modified -site 277..279
FT /label= N-glycos site
FT Modified -site 349..351
FT /label= N-glycos site
FT Modified -site 600..602
FT /label= N-glycos _site
PN W09100292-A.

PD 10-JAN-1991.
PF 22-JUN-1990; U03586.
PR 23-JUN-1989; US-370673.
PA (GETH) GENENTECH INC.
PI Chang M, Goeddel D, Lowe D;
DR WPI; 91-036711/05.
DR N-PSDB; Q10324.
PT Natriuretic protein receptor B - for diagnosis and treatment of
PT kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS Claim 3; Fig 1; 49pp; English.
CC The sequence was derived from the DNA encoding natriuretic peptide
CC receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC kinase activity. The DNA can be inserted into expression vectors
CC for the prodn. of the protein, opt. after being mutated to produce
CC NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
CC 114,952). The protein (or variants) can be used in treatment of
CC natriuretic peptide disorders, and also to isolate peptides using
CC affinity chromatography. Antibodies with affinity for NPRB can
CC also be prepd.
SQ Sequence 1047 BP; 87 A; 15 C; 83 G; 51 T;

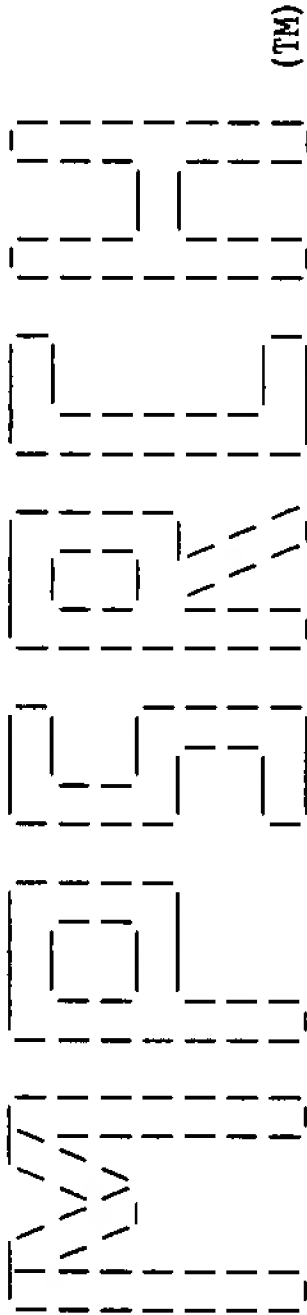
Query Match 3.4%; Score 94; DB 2; Length 1047;
Best Local Similarity 8.8%; Pred. No. 4.63e-38;
Matches 84; Conservative 267; Mismatches 594; Indels 7; Gaps 7;

Db 84 avdnknyhdnngnvcvynaasvarnashwrnnnnntagavasgnsakndhytrnvtg 143

May 31 09:20

US-08-426-509-4.rpr

1



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri May 31 09:28:01 1996; MasPar time 27.32 Seconds
624.277 Million cell updates/sec

Tabular output not generated.

Title: >US-08-426-509-4
Description: (1-675) from US08426509.pep
Perfect Score: 4998
Sequence: 1 MDTKSIIEELLKRSQQKKK.....RPTFQQLSSIEPLREKDKH 675

Scoring table: PAM 150
Gap 11

Searched: 82306 seqs, 25270970 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir46

1:ann1 2:ann2 3:ann3 4:unann1 5:unann2 6:unann3 7:unann4
8:unann5 9:unann6 10:unann7 11:unann8 12:unrev1
13:unrev2

Statistics: Mean 49.410; Variance 111.531; scale 0.443

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	2059	41.2	659	11	S28912 protein-tyrosine kin	0.00e+00
2	2031	40.6	660	11	JN0471 protein-tyrosine kin	0.00e+00
3	2016	40.3	659	11	B45184 B cell progenitor ki	0.00e+00
4	1978	39.6	630	11	JU0228 protein tyrosine kin	0.00e+00
5	1791	35.8	620	11	S33253 protein-tyrosine kin	2.24e-299
6	1762	35.3	527	13	S13763 protein-tyrosine kin	6.06e-294
7	1761	35.2	619	11	A47333 T-cell-specific tyro	9.33e-294
8	1747	35.0	608	11	JU0227 protein-tyrosine kin	3.90e-291
9	1745	34.9	527	11	A55631 protein-tyrosine kin	9.23e-291
10	1736	34.7	619	11	JN0472 protein-tyrosine kin	4.47e-289
11	1719	34.4	625	11	A43030 protein-tyrosine kin	6.79e-286
12	1686	33.7	602	11	JU0215 tyrosine kinase, tec	1.01e-279

May 31 09:20

US-08-426-509-4.rpr

2

13	1532	30.7	590	1	TVFFDS	protein-tyrosine kin	5.84e-251
14	1138	22.8	506	4	S24553	protein-tyrosine kin	7.57e-178
15	1093	21.9	981	1	FOMVGM	gag-abl polyprotein	1.48e-169
16	1090	21.8	557	10	A00629	protein-tyrosine kin	5.29e-169
17	1090	21.8	1130	4	S08519	protein-tyrosine kin	5.29e-169
18	1085	21.7	697	7	A26132	gag-abl-pol polyprot	4.41e-168
19	1087	21.7	1123	4	A39962	kinase-related trans	1.89e-168
20	1074	21.5	1130	1	TVHUA	protein-tyrosine kin	4.66e-166
21	1068	21.4	505	4	S24550	protein-tyrosine kin	5.92e-165
22	1068	21.4	507	4	A39939	protein-tyrosine kin	5.92e-165
23	1070	21.4	1520	1	TVFFFA	protein-tyrosine kin	2.54e-165
24	1053	21.1	509	1	OKHULK	protein-tyrosine kin	3.39e-162
25	1055	21.1	536	10	S33569	protein-tyrosine kin	1.45e-162
26	1055	21.1	536	4	S29626	protein kinase yrk (1.45e-162
27	1052	21.0	362	4	S24551	protein-tyrosine kin	5.17e-162
28	1048	21.0	509	4	A23639	protein-tyrosine kin	2.81e-161
29	1051	21.0	537	1	TVHUSY	protein-tyrosine kin	7.90e-162
30	1047	20.9	537	4	A43806	protein-tyrosine kin	4.29e-161
31	1046	20.9	537	1	TVHUSR	protein-tyrosine kin	6.55e-161
32	1045	20.9	1146	4	B35962	protein-tyrosine kin	1.00e-160
33	1045	20.9	1182	4	A35962	protein-tyrosine kin	1.00e-160
34	1041	20.8	526	1	TVFV60	protein-tyrosine kin	5.43e-160
35	1038	20.8	533	1	TVCHS	protein-tyrosine kin	1.93e-159
36	1041	20.8	541	4	A43610	protein-tyrosine kin	5.43e-160
37	1041	20.8	542	1	TVHUSC	protein-tyrosine kin	5.43e-160
38	1039	20.8	557	1	TVFVS2	protein-tyrosine kin	1.27e-159
39	1040	20.8	568	1	TVFVS1	protein-tyrosine kin	8.29e-160
40	1039	20.8	587	1	TVFVPR	protein-tyrosine kin	1.27e-159
41	1036	20.7	532	4	A34104	protein-tyrosine kin	4.50e-159
42	1036	20.7	532	4	B34104	protein-tyrosine kin	4.50e-159
43	1031	20.6	505	1	TVHUHC	protein-tyrosine kin	3.73e-158
44	1030	20.6	523	1	TVFVMT	protein-tyrosine kin	5.69e-158
45	1029	20.6	545	7	S52313	pp62v protein - Rous	8.68e-158

ALIGNMENTS

```
RESULT 1
ENTRY S28912 #type complete
TITLE protein-tyrosine kinase (EC 2.7.1.112) atk - human
ORGANISM #formal name Homo sapiens #common name man
DATE 25-Feb-1994 #sequence_revision 01-Sep-1995 #text_change
01-Sep-1995

ACCESSIONS S28912
REFERENCE S28912
#authors Vetrie, D.; Vorechovsky, I.; Sideras, P.; Holland, J.;
Davies, A.; Flinter, F.; Hammarstroem, L.; Kinnon, C.;
Levinsky, R.; Bobrow, M.; Smith, C.I.E.; Bentley, D.R.
#journal Nature (1993) 361:226-233
#title The gene involved in X-linked agammaglobulinaemia is a member
of the src family of protein-tyrosine kinases.
#accession S28912
##status preliminary
##molecule_type mRNA
##residues 1-659 ##label VET
KEYWORDS phosphotransferase
SUMMARY #length 659 #molecular-weight 76281 #checksum 9489

Query Match 41.2%; Score 2059; DB 11; Length 659;
Best Local Similarity - 48.9%; Pred. No. 0.00e+00;
Matches 338; Conservative 144; Mismatches 149; Indels 60; Gaps 28;

Db 5 ilesiflkrsgqkktsplnfkkrflflltvhklsvyeydferrgrgskgsidvekitcv 64
||| :||||||||| || :| ||| : ||||| : |||:||||: || ||
```


May 31 09:20

US-08-426-509-4 rpr

5

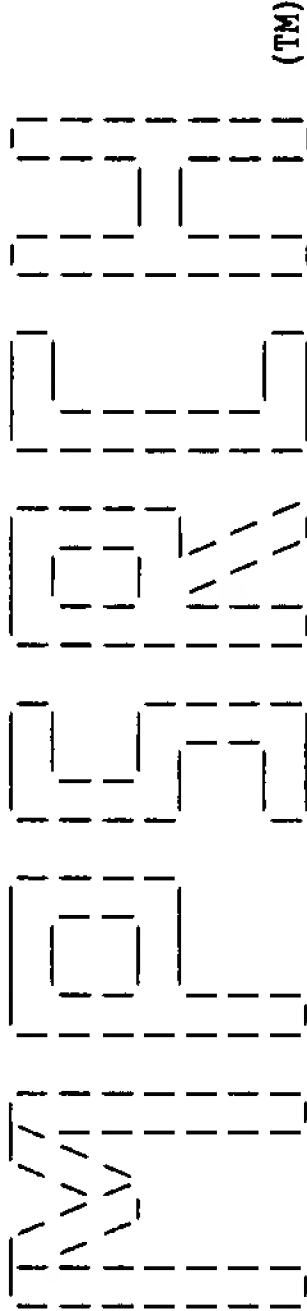
```
RESULT      3
ENTRY
TITLE      B45184      #type complete
           B cell progenitor kinase, BPK=cytoplasmic tyrosine kinase -
           mouse
ORGANISM    #formal name Mus musculus #common name house mouse
DATE        30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change
           12-May-1995
ACCESSIONS  B45184
REFERENCE   A45184
#authors    Tsukada, S.; Saffran, D.C.; Rawlings, D.J.; Parolini, O.;
           Allen, R.C.; Klisak, I.; Sparkes, R.S.; Kubagawa, H.;
           Mohandas, T.; Quan, S.; Belmont, J.W.; Cooper, M.D.;
           Conley, M.E.; Witte, O.N.
#journal    Cell (1993) 72:279-290
#title      Deficient expression of a B cell cytoplasmic tyrosine kinase
           in human X-linked agammaglobulinemia.
#cross-references MUID:93145329
#accession  B45184
##status    preliminary; not compared with conceptual translation
##molecule_type nucleic acid
##residues   1-659 ##label TSU
##cross-references NCBI:P123834
##experimental_source 70z/3 pre-B cell lineli
##note       sequence extracted from NCBI backbone
CLASSIFICATION #superfamily SH3 homology; protein kinase homology; SH2
               homology
FEATURE
221-269      #domain SH3 homology #label SH3\
281-377      #domain SH2 homology #label SH2\
400-658      #domain protein kinase homology #label KIN
SUMMARY      #length 659 #molecular-weight 76326 #checksum 9917
Query Match      40.3%; Score 2016; DB 11; Length 659;
Best Local Similarity 48.0%; Pred. No. 0.00e+00;
Matches 332; Conservative 145; Mismatches 154; Indels 60; Gaps 29;
Db 5 ilesiflkrsgqkktspfnfkrflfltvhklsvyeydferrgrsgkksldvekitcv 64
   |||::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 6 ILEELLKRSQQKKKMSPNNYKERLFVLTKTNLSYYEYD--KMKRGSRKGSIEIKIRCV 63
Db 65 etvipeknppperqiprrgeesemesegisierfypfvvydegplyvfsepteelrkpw 124
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 64 EKV-----L-----E-EQTPV-ER-QYPFQIVYKDGCLLYVYASNEESRSQW 102
Db 125 ihqlknvirynsdlvqkyhpcfidgqylccsqtaknamgcqilen-rn--gs1---k-- 176
   : | : | | | : | | | : | | | : | | | : | | | : | | | : | | : |
Qy 103 LKALQKEIRGNPHLLVKYHSGFFVDGKFLCCQSQCKAAPGCTLWEAYANLHTAVNEKHR 162
Db 177 -pgsshr--ktkk--p-l----pp--tp--ee-dqllkklppteapis-tt-el--kk 218
   | : | | | : | | | : | | | : | | | : | | | : | | | : | | | : |
Qy 163 VPTFPDRVLKIPRAVPVLKMDAPSSSTTLAQYDNESKKNYGSQPPSSSTSLAQYDSNKK 222
Db 219 v-val--yd--ympmna-ndl-qlrkgeeyfileesnlpwward-kn--gqegy-ipsn 267
   : : : | | : | | : | | | | | | | | | | | | | | | | | | | | |
Qy 223 IYGSQPNFMQYIPREDFDWWQVRKLSSSSSEDEVASSNQKERNVNHHTTSKISWEFPES 282
Db 268 yiteaedsiemewyskhtsrsgaeqlkqegkgffivrdsskagkytvsfvakstgep 327
   : | | : : | | : : : | | | | | | | | | | | | | | | | | | | |
Qy 283 SSSEEEENLDDYDWFAGNISRSQSEQLLRQKKEGAFMVRNSSQVGMVTVSLFSKAVNDK 342
Db 328 qgvirhyvvcstpqsgvylaekhlfstipelinhyqhnsaglisrlkypvskqknapst 387
   | : | | | : : : | | | : | | | | | | | | | | | | | | | | | | |
```

May 31 09:20

US-08-426-509-4 rpr

6

```
Qy 343 KGTVKHYHVHTNAENKLYLAENYCFDSIPKLIHYHQHNSACMITRLRHPVSTFANKVPDS 402
Db 388 aglgygsweidpdkdltflkelgtqgfgvvkygkwrgqydvaiкмiregmsedeefieeak 447
   : | | | | : : : | | | | : | | | | | | | | | | | | | | | | | | |
Qy 403 VSLNGIWEIKREEITLLKELGSCQFCVVQLCKWKQGYDVAVKMIKEGSMSEDEFFQEAQ 462
Db 448 vmmnlsheklvglygvctkqrpfiiteyemangcllnylremhrfrfqtqqllemckdvce 507
   || : | | | | | | | | : | | : | | | | : | | | | | | | | | | | |
Qy 463 TMWKLSPKLVKFEYCVCSKEYPIYIVTEYISNGCLLNYLRSHGKLGLEPSQLLEMCYDVCE 522
Db 508 ameyleskqflhrdlaarnclvndqgvkvslpglsryviddeytsvsgskfpvrwspe 567
   : | : | | | | | | | | | | : | | | | : | | | | | | | | | | | |
Qy 523 GMAFLESHQFIHRDLAARNCLVDRDLGVKVSDFGMYRYVLDQYVSSVGTKEFPVKWSAPE 582
Db 568 vlmyskfssksdiwafgvlmweiyslgkmpyerfnsetaehiaqglrlyrphlaservy 627
   | : | | | | | | | | | | : | | | | | | | | : | | : | | | | | | : |
Qy 583 VFHYKYSKSDVWAFGILMWEVFSLGKQPYDLYDNSQVVLKVSQGHRLYRPHLASDTIY 642
Db 628 timyschwhekaderpskillsnildvmddee 658
   | | | | | | | : | | : | | | | | | | : : :
Qy 643 QIMYSCHWHELPEKRPTFQQLLSSIEPLREKD 673
RESULT      4
ENTRY      JU0228      #type complete
TITLE      protein tyrosine kinase (EC 2.7.1.-) tec IV - mouse
ORGANISM    #formal name Mus musculus #common name house mouse
DATE        30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
           07-Jul-1995
ACCESSIONS  JU0228
REFERENCE   JU0228
#authors    Mano, H.; Sato, K.; Yazaki, Y.; Hirai, H.
#submission submitted to JIPID, April 1993
#accession  JU0228
##molecule_type DNA
##residues   1-630 ##label MAN
##experimental_source myeloid
CLASSIFICATION #superfamily protein kinase homology; SH2 homology; SH3
               homology
KEYWORDS     kinase-related transforming protein
FEATURE
185-233      #domain SH3 homology #label SH3\
246-344      #domain SH2 homology #label SH2\
367-625      #domain protein kinase homology #label KIN\
375-383      #region protein kinase ATP-binding motif\
515          #active_site Asp (aspartylphosphate intermediate)
           #status predicted
SUMMARY      #length 630 #molecular-weight 73651 #checksum 4354
Query Match      39.6%; Score 1978; DB 11; Length 630;
Best Local Similarity 46.7%; Pred. No. 0.00e+00;
Matches 315; Conservative 141; Mismatches 164; Indels 54; Gaps 28;
Db 1 mnfntileeilkrsgqkktksllnykerlcvlpksvlsyve-graekkyrkvidiski 59
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1 MDTKSIIEELLKRSQKKKMSPNNYKERLFVLTKTNLSYYEYDKMKRGSRKGSIEIKKI 60
Db 60 kcveivknndgvpicqmkfpfvvhdantliyfapspqsrdrwvklkeeknnnnmik 119
   : | | | | : : | : | | | | | | | | | | | | | | | | | | | | | |
Qy 61 RCVEKVNLEEQT-PVERQYFPQIVYKDGCLLYVYASNEESRSQWLKALQKEIRGNPHLLVK 119
Db 120 yhpkwadgsgyccqrcteklapgc---eky-nlfercir-k-tlpape-i-kkrrpppp 171
   | | : | | : | | | | | | | | | | | | | | | | | | | | | | |
Qy 120 YHSGFFVDGKFLCCQSQCKAAPGCTLWEAYANLHTAVNEEKHRVPTFPDRVLKIPRAVPV 179
```



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Fri May 31 10:49:54 1996; MasPar time 1524.93 Seconds
1160.661 Million cell updates/sec

Tabular output not generated.

Title: >US-08-426-509-3
Description: (1-2500) from US08426509.seq
Perfect Score: 2500
N.A. Sequence: 1 CCGCTTTTGTCTAGAGCTT.....AAAAA.....AAAAA 2500
Comp: GCGCAAAACGAATCTCGAA.....TTTTTTTTTTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 264399 seqs, 353985056 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embl-new11
1:BCT 2:FUN 3:INV1 4:INV2 5:INV3 6:MAM 7:ORG 8:PLN
9:PRI1 10:PRI2 11:PRI3 12:PRO1 13:PRO2 14:ROD 15:SYN
16:UNC 17:VRT 18:VIR
Database: genbank91
19:BCT1 20:BCT2 21:BCT3 22:BCT4 23:BCT5 24:BCT6 25:BCT7
26:INV1 27:INV2 28:INV3 29:INV4 30:INV5 31:MAM1 32:MAM2
33:PAT1 34:PAT2 35:PAT3 36:PHG 37:PLN1 38:PLN2 39:PLN3
40:PLN4 41:PLN5 42:PLN6 43:PLN7 44:PRI1 45:PRI2 46:PRI3
47:PRI4 48:PRI5 49:PRI6 50:PRI7 51:PRI8 52:PRI9 53:ROD1
54:ROD2 55:ROD3 56:ROD4 57:ROD5 58:ROD6 59:ROD7 60:STR
61:SYN 62:UNA 63:VRL1 64:VRL2 65:VRL3 66:VRL4 67:VRL5
68:VRL6 69:VRT1 70:VRT2 71:VRT3
Database: genbank-new11
72:BCT1 73:BCT2 74:INV1 75:INV2 76:MAM 77:PHG 78:PLN
79:PRI1 80:PRI2 81:PRI3 82:ROD 83:STR 84:SYN 85:UNA
86:VRL 87:VRT
Database: u-embl44 91
88:part1

Statistics: Mean 12.481; Variance 6.090; scale 2.049

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description	Pred. No.
1	2397	95.9	2456	44	HSBMXGENE	H.sapiens Bmx mRNA fo	0.00e+00
2	329	13.2	2468	55	MUSBTKCDNA	Mouse Bruton agammagl	6.86e-243
3	327	13.1	2485	55	MUSBPK	Mus musculus B cell c	3.56e-241
4	321	12.8	2560	44	HSATK	H.sapiens atk mRNA fo	4.95e-236
5	312	12.5	2546	55	MUSEMBX	Mus musculus mRNA seq	2.53e-228
6	259	10.4	2564	52	HUMTYRKINA	Human tyrosine kinase	3.91e-183
7	247	9.9	2574	59	S53716	Tec=protein tyrosine	5.85e-173
8	246	9.8	2578	54	MWTECMR	Mouse mRNA for protei	4.11e-172
9	232	9.3	1680	54	MMU19607	Mus musculus tyrosine	2.83e-160
10	232	9.3	2204	82	MMU16145	Mus musculus tyrosine	2.83e-160
11	232	9.3	2204	54	MMU16145	Mus musculus tyrosine	2.83e-160
12	232	9.3	2204	14	MM16145	Mus musculus tyrosine	2.83e-160
13	232	9.3	2221	57	MUSRLK	Mus cookii resting ly	2.83e-160
14	232	9.3	2342	57	MUSPTKRL18	Mouse mRNA for protei	2.83e-160
15	215	8.6	4224	57	MUSTLK	Mouse tlk mRNA for ty	5.91e-146
16	215	8.6	4231	57	MUSTYRKIN	Mouse tyrosine kinase	5.91e-146
17	215	8.6	4294	56	MUSITKA	Mouse T-cell-specific	5.91e-146
18	209	8.4	2480	55	MUSEMTX	Mus musculus mRNA seq	6.47e-141
19	200	8.0	3650	51	HUMPTKA	Human mRNA for Tec pr	2.25e-133
20	201	8.0	4366	50	HUMLYK	Human mRNA for LYK, c	3.28e-134
21	201	8.0	6383	52	HUMTKTCS	Homo sapiens T cell-s	3.28e-134
22	175	7.0	192	47	HSU08341	Human clone NTK38 tyr	1.52e-112
23	137	5.5	2940	28	DROSRC28C	D.melanogaster src-re	2.52e-81
24	97	3.9	1253	30	SLSRK2	S.lacustris srk2 mRNA	2.17e-49
25	96	3.8	1190	30	SLSRK3	S.lacustris srk3 mRNA	1.31e-48
26	94	3.8	2091	63	FCSH22A	Hardy-Zuckerman 2 fel	4.71e-47
27	92	3.7	1702	30	SLSRK4	S.lacustris srk4 mRNA	1.68e-45
28	90	3.6	1829	30	SLSRK1	S.lacustris srk1 mRNA	5.91e-44
29	90	3.6	2200	58	RNFGR	R.norvegicus FGR mRNA	5.91e-44
30	87	3.5	2526	71	XHCFYN	Xiphophorus c-fyn (Xf	1.21e-41
31	85	3.4	399	27	DMSRC4	Drosophila melanogast	4.13e-40
32	82	3.3	2133	53	MMCFGRMR	Murine c-fgr mRNA.	8.03e-38
33	82	3.3	2179	53	MMCFGR	Murine c-fgr53 proto-	8.03e-38
34	82	3.3	2241	63	ALRPP60A	Rous sarcoma virus ph	8.03e-38
35	83	3.3	3653	55	MUSCABL	Mouse tyrosine kinase	1.39e-38
36	83	3.3	3884	68	REABMLVA	Abelson (P160) murine	1.39e-38
37	82	3.3	4413	63	ALSRCAC	Rous sarcoma virus de	8.03e-38
38	82	3.3	4503	63	ASVPR225T	Avian Sarcoma Virus P	8.03e-38
39	83	3.3	5058	55	MUSABLT5	Mouse testis-specific	1.39e-38
40	82	3.3	5188	63	ACSCSRC	Avian sarcoma virus p	8.03e-38
41	83	3.3	5893	68	REAMLV	Abelson murine leukem	1.39e-38
42	83	3.3	5894	68	MLAPRO	Abelson murine leukem	1.39e-38
43	79	3.2	208	57	MUSPTKG	Mouse protein-tyrosin	1.51e-35
44	80	3.2	1759	69	CHKSRC	Chicken c-src gene, c	2.65e-36
45	78	3.1	1759	70	GGCSRC	Gallus gallus gene c-	8.61e-35

ALIGNMENTS

RESULT	1	HSBMXGENE	2456 bp	RNA	PRI	22-AUG-1995
LOCUS		H.sapiens Bmx mRNA for cytoplasmic tyrosine kinase.				
DEFINITION		H.sapiens Bmx mRNA for cytoplasmic tyrosine kinase.				
ACCESSION		X83107				
KEYWORDS		cytoplasmic; Tyrosine kinase.				
SOURCE		human.				
ORGANISM		Homo sapiens				
		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;				
		Vertebrata; Osteichthyes; Sarcopterygii; Mammalia; Eutheria;				

Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2456)
AUTHORS Tamagnone, L., Lahtinen, I., Mustonen, T., Virtaneva, K., Francis, F., Muscatelli, F., Alitalo, R., Smith, C. I., Larsson, C. and Alitalo, K.
TITLE BMX, a novel nonreceptor tyrosine kinase gene of the BTK/ITK/TEC/TXK family located in chromosome Xp22.2
JOURNAL Oncogene 9 (12), 3683-3688 (1994)
MEDLINE 95460827
REFERENCE 2 (bases 1 to 2456)
AUTHORS Tamagnone, L.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-1994) to the EMBL/GenBank/DBJ databases. L. Tamagnone, University of Helsinki, Molecular/Cancer Biology Lab., PL21 (Haartmaninkatu 3), 00014 Helsinki, FINLAND
COMMENT Related sequence: U08341.

NCBI gi: 951234	Location/Qualifiers
1..2456	
/organism="Homo sapiens"	
/tissue type="bone marrow"	
/clone_lib="endothelial cell cdna"	
/chromosome="X"	
/map="p22.2"	
34..2061	
/gene="bmx"	
/note="pid:el28520; NCBI gi: 951235"	
/codon start=1	
/translation="MDTKSILELLLRSQKKMSPNNYKERLFVLTKNLSYYEYD KMKRGSRKGSIEIKKIRCVKVNLEEQTPVERQYFPQIVYKDGLLYVYASNEESRSQW LKALQKEIRGNPHLLVYHSGFFVDGKFLCCQQCKAAPGCTLWEAYANLHTAVNEEK HRVPTFPDRVLIKIPRAVPVLKMDAPSSSTTLAQYDNEKKNYGSQPPSSSTSIAQYDS NSKKIYGSQPNFMQYIPREDFPDWQVRKLLKSSSSESDVASNQKERNVNHHTTSKIS WEFPSSSSEEEENLDDYDFWAGNISRSQSEQLLRQKREGAFVYRNSQVGMYYTSL FSKAVNDKKGTVKHYHVHTNAENKLYLAENYCFDSIPKLIHYHQHNSAGMITRLRPV STKANKVPDSVSLNGIWEIKREEITLLKELGSGQFGVVQLGKWKQGYDVAVVKMIKEG SMSEDEFFQEAQTMKLSHPKLVKFYGVCSKEYPYIIVTEYISNGCLLNYLRSHGKGL EPSQILLEMCYDVCEGMALFESHQFIHRDLAARNCLVDRDLCKVSDFGMTRVYLDQY VSSVGTFPPVKSAPFHYFYKYSKSDVWAFGILMWEVFSLGKQPYDLYDNSQVVLK VSQGHRLYRPHLASDTIYQIMYSCWHELPKRPFTFQQLLSSIEPLREKOKH"	
BASE COUNT 805 a 495 c 549 g 607 t	
ORIGIN	
Query Match 95.9%; Score 2397; DB 44; Length 2456; Best Local Similarity 99.4%; Pred. No. 0.00e+00; Matches 2416; Conservative 0; Mismatches 13; Indels 1; Gaps 1;	
Db 24 ggaatgataatggatacaaaatctattctagaagaactcttctcaaaagatcacagca 83 	
Qy 72 GGATGATAATGGATACAAAATCTATTCTAGAGAACTTCTTCTCAAAAGATCACAGCA 131 	
Db 84 aaagaagaaatgtcaccaataattacaaagaacggcttttgttttgaccacaaacaa 143 	
Qy 132 AAAGAAGAAATGTCAACAAATAATTACAAAGACGGCTTTTGTGTTTGACCAACAAA 191 	
Db 144 cctttcctactatgaatgacaaaatgaaaaggggcagacagaaaaggatccattgaaat 203 	
Qy 192 CCTTTCCTACTATGAATATGACAAAATGAAAAGCGGCATTTTGTGTTTGACCAACAAA 251 	
Db 204 taagaaaatcagatgtgtggaaaagtaaaatctcgaggagcagacgcctgtagagagaca 263 	
Qy 252 TAAGAAAATCAGATGTCTGGAGAAAGTAAATCTCGAGGACGACGCGCCTGTAGAGAGACA 311 	
Db 264 gtaccatttcagattgtctataaagatgggcttctctatgtctatgcatcaaatgaaga 323	

Qy 312 GTACCCATTTTCAGATTGCTATAAAGATGGGCTTCTCTATGCTCATCAATGAACA 371 	
Db 324 gagccgaagtcagtggtgaaagcattacaaaaagagataaggggtaaccccccacctgct 383 	
Qy 372 GAGCCGAAGTCAGTGGTTGAAAGCATTTACAAAAGACAGATAAGGGGTAAACCCACCTGCT 431 	
Db 384 ggtcaagtaccatagtggtgtcttcgtggacgggaagttcctgtgttgcagcagagctg 443 	
Qy 432 GGTCAAGTACCATAGTGGGTTCTTCGTGGACGGGAAGTTCCCTGTGTGTCACGACGAGCTG 491 	
Db 444 taaagcagcccgaggatgtaccctctgtgggaagcatatgctaactgcatactgcagtcaa 503 	
Qy 492 TAAAGCAGCCCCAGGATGTACCCCTCTGGGAAGCATATGCTAAATCTGCATACTGCAGTCAA 551 	
Db 504 tgaagagaaaacacagagttcccaccctccacagacagagtgctgaagataacctcgggcagt 563 	
Qy 552 TGAAGAGAAAACACAGAGTCCCACCTTCCCGACAGACAGTGCCTGAAGATACCTCGGGCAGT 611 	
Db 564 tcctgttctcaaaatggatgcaccatcttcaagtaccactctagcccaaatatgacaacga 623 	
Qy 612 TCCTGTTCTCAAAATGGATGCACCATCTTTCAAAGTACCACCTTAGCCCAATATGACAACGA 671 	
Db 624 atcaaaagaaaaactatggtcccagccaccatcttcaagtaccagtcagtcagcgaatatga 683 	
Qy 672 ATCAAAAGAAAATACTATGCTCCAGCCACCATCTTCAAGTACCAGTCTAGCGCAATATGA 731 	
Db 684 cagcaactcaagaataatctatggtcccagcccaacttcaacatgcagtatattccaag 743 	
Qy 732 CAGCAACTCAAGAAAATCTATGGCTCCCAGGCCAAACTTCAACATGCAGTATATTCCAAG 791 	
Db 744 ggaagacttccctgactggtggcaagtgaagaaaaactgaaaagtgcagcagcagtggaaga 803 	
Qy 792 GGAAGACTTCCCTGACTGCTGGCAAGTAAAGAAAAGTAAAGTAGCCAGCAGCAGTGAAGA 851 	
Db 804 tgttgaagcagtaaccacaaagaaagaatgtgaatcacaccacctcaagatttcatg 863 	
Qy 852 TGTTCGAGCAGTAACCAAAAAGAAAGAAATGTGAATCACACCACCTCAAGATTTTCATG 911 	
Db 864 ggaattccctgagtcagtcagtcctgaagaagaggaacacctggatgattatgactggtt 923 	
Qy 912 GGAAATTCCTGAGTCAAGTTTCATCTGAAGAAGAGGAAAACCTGGATGATTAATGACTGGTT 971 	
Db 924 tgctgttaacatctccagatcacaaatctgaacagtggaatgtacacagtgctccttattagtaa 983 	
Qy 972 TGCTGGTAACATCTCCAGATCACAAATCTGAACAGTTACTCAGACAAAAGGGGAAAAGAGG 1031 	
Db 984 agcatttatggttagaaattcgagcccaagtgggaatgtacacagtgctccttattagtaa 1043 	
Qy 1032 AGCATTTATGGTTAGAAATTCGAGCCCAAGTGGGAATGTACACAGTGTCTTATTAGTAA 1091 	
Db 1044 ggctgtgaatgataaaaaaggaactgtcaaacattaccagtgcatacacaaatgctgagaa 1103 	
Qy 1092 GGCTGTGAATGATAAAAAGGAACGTGTCAAAACATTACCAGCTGCATACAAATGCTGAGAA 1151 	
Db 1104 caaattatacctggcagaaaaactactgttttgattccattccaaagcttattcattatca 1163 	
Qy 1152 CAAATTATACCTGGCAGAAAACACTGTGTTTGAITCCATTCCAAAGCTTATTATATCA 1211 	
Db 1164 tcaacacaaattcagcaggcatgatcacacggctcgcgccacctgtgtcaacaaaggccaa 1223 	
Qy 1212 TCAACACAAATTCAGCAGGCATGATCACACGGCTCGCGCACCTGTGTCAACAAAGGCCAA 1271 	
Db 1224 caaggtccccgactctgtgtccctgggaatggaatctgggaactgaaaagagagagat 1283	

WKKZ

Qy 1272 CAAGGTCCTCCGACTCTGTCCTCCCTGGGAATGGAATCTGGAACTGAAAAGAGAAGAT 1331
Db 1284 taccttggtgaaggagctgggaagtggcaggttggaagtgtccagctggcaagtggaa 1343
Qy 1332 TACCTTGCTGAAGGAGCTGGGAAGTGGCAGTTTGGAGTGGTCCAGCTGGGCAAGTGAA 1391
Db 1344 ggggcagtatgatgttgctgttaagatgatcaaggagggtccatctcagaaagtgaatt 1403
Qy 1392 GGGGCAGTATGATGTTGCTGTTAAGATGATCAAGGAGGCTCCATCTCAGAAGATGAATT 1451
Db 1404 ctttcaggaggccagactatgatgaaactcagccatcccaagctgggttaattctatgg 1463
Qy 1452 CTTTCAGGAGGCCAGACTATGATGAAACTCAGGCATCCCAAGCTGGTAAATTCTATGG 1511
Db 1464 agtgtgtcaaggaaataccccatatacatagtactgaatatataagcaatggctgctt 1523
Qy 1512 AGTGTGTTCAAGGAATACCCCATATACATAGTACTGACTGAATATATAAGCAATGGTCTT 1571
Db 1524 gctgaattacctgaggagtcacggaagagcttgaaacctcccagctctcttagaaatgtg 1583
Qy 1572 GCTGAATTACCTGAGGAGTCACGGAAGGACTTGAACCTTCCCAGCTCTTAGAAATGTG 1631
Db 1584 ctacgatgtctgtgaaggcatggcctcttggaagatcaccaattcatacaccgggactt 1643
Qy 1632 CTACGATGCTGTGAAGCATGGCTTCTTGGAGAGTCACCAATTATACACCCGGGACTT 1691
Db 1644 ggctgtcgttaactgcttggtggacagadatctctgtgtgaagtatctgactttggaat 1703
Qy 1692 GCCTGCTCGTAACTGCTTGGTGCACAGAGATCTCTGTGTGCAAGTATCTGACTTTGCAAT 1751
Db 1704 gacaaggtatgttcttgatgaccagtatgtcagttcagtcagtcaggaacaaagttccagtc 1763
Qy 1752 GACAAGGTATGTTCTTGATGACCGATGCTCAGTCTCAGTCGGGAACAAGTTCCAGTCAA 1811
Db 1764 gtggtcagctccagaggtgtttcattactcaatacacagcagcaagtccagtgatgggc 1823
Qy 1812 GTGCTAGCTCCAGAGGTGTTTCATTACTTCAATATACAGCAGCAATCAGACGTATGGGC 1871
Db 1824 attgggatctgatgtgggaggtgttcagcctgggaagcagccctatgacttgtatga 1883
Qy 1872 ATTTGGATCTCTGATGTGGAGCTGTTTCCAGCTGGGAACAGCCCTATGACTTGTATGA 1931
Db 1884 caactccagtggttctgaaggtctccaggggccacaggctttaccggccccacctggc 1943
Qy 1932 CAACTCCAGGTGGTCTTCAAGTCTCCAGGGCCACAGSGCTTTACCGGCCCCACCTGGC 1991
Db 1944 atcggacaccatctaccagatcatgtacagctgctggcagagcttccagaaaagcgctc 2003
Qy 1992 ATCGGACACCATCTACCAGATCATGTACAGCTGCTGGCAGCAGCTTCCAGAAAAGCGTCC 2051
Db 2004 cacatttcagaactcctgtcttccattgaaccacttcgggaaaaagacaagcattgaag 2063
Qy 2052 CACATTTCAGCAACTCCTGTCTCTCCATTGAACCACTTCGGGAAAAAGACAAGCATTGAAG 2111
Db 2064 aagaaattaggagtgtgataaagaatgaatagatgtggtggccagcatttctcattcattt 2123
Qy 2112 AAGAAATTAGGAGTGTCTATAGCAATGAATATAGATGCTGGCCACCATTTTCATTCAATT 2171
Db 2124 taaggaaagtgaaggacataagtaatttagctagtttttaataagtgttctctgtattg 2183
Qy 2172 TAAGGAAGTAGCAAGCCATAATCTAAATTAGCTAGTTTTTAATAGTGTCTCTGTATTG 2231
Db 2184 tctattattagaataatgaacaaggcagggaacaaaagattcccttgaaatttagatcaaa 2243

Qy 2232 TCATTATTAGAAATGAACAAGCAGGAACAACAACATTCCCTTGAAATTTAGGTCAAA 2291
Db 2244 ttagtaatttgttttatgtctgctctctgatatataaacacttccagcctatageagaageac 2303
Qy 2292 TTAGTAATTTTGTTT-ATGCTGCCCTGATATAACACTTTCCAGCCTATAGCACAAGCAC 2350
Db 2304 attttcagactgcaatatagagactgtgttcattgtgttaaagactgacgagaactgaaaaa 2363
Qy 2351 ATTTTCAGACTGCAATATAGAGACTGTGTTCAATGTCTAAAGACTGAGCAGAACTGAAAAA 2410
Db 2364 ttacttattgatatattcattcttttctttatattgtcattgtcacacaattaaatatac 2423
Qy 2411 TTACTTATTGGATATTCAATCTTTTCTTTATATTGTCATTGTCACAACAATTAAATATAC 2470
Db 2424 tacciaagtacagaaatgtggaaaaa 2453
Qy 2471 TACCAAGTACAAAAA 2500

RESULT 2
LOCUS MUSBTKCDNA 2468 bp mRNA ROD 18-MAY-1995
DEFINITION Mouse Bruton agammaglobulinemia tyrosine kinase (Btk) mRNA,
complete cds.
ACCESSION L29788
KEYWORDS Bruton agammaglobulinemia tyrosine kinase.
SOURCE Mus musculus cDNA to mRNA.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Glires; Rodentia;
Sciurognathi; Myomorpha; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2468)
AUTHORS Sideras,P., Muller,S., Shields,H., Jin,H., Khan,W.N., Nilsson,L.,
Parkinson,E., Thomas,J.D., Branden,L., Iarsson,I., Paul,W.E.,
Rosen,F.S., Alt,F.W., Vetrie,D., Smith,C.I.E. and Xanthopoulos,K.G.
TITLE Genomic organization of mouse and human Bruton's agammaglobulinemia
tyrosine kinase (Btk) loci
JOURNAL J. Immunol. 153 (12), 5607-5617 (1994)
MEDLINE 95081608
COMMENT NCBI gi: 625143
FEATURES Location/Qualifiers
source 1..2468
/organism="Mus musculus"
/germline
/sequenced_mol="cDNA to mRNA"
133..2112
/gene="Btk"
/map="X"
/note="NCBI gi: 625144"
/codon_start=1
/product="Bruton agammaglobulinemia tyrosine kinase"
/translation="MAAVILESIFLRSQKKKTSPLNFKRILFLTFVHKLSYYEYDF
ERGRGSKKSIDVEKITCVETVIPEKNPPPERQIPRRGESSEMEQISIIERFPYPF
QVYDEGPLYVFSPTTELKRWIHQLNVIYNSDLVQKYPHPCFIDQYLCGSQTAK
NAMCQIENRNGSLKPGSSHRKTKPLPTPEEQILKKPLPPEPTAAPISTTELKK
VVALYDYPMNANDLQLRKGEYFILEESNLPWWRARDKNQEGYIFSNYITEAEDSI
EMEWYSKHMTRSQAQELLKQEGKEGFIVRDSSKAGKYTVSVFAKSTGEPQGVIRHY
VVCSTPOSQYLLAEKHLFSTIPELINYHOHNSAGLISRLKYPVSKONKNAPSSTAGLGY
GSWEIDPKDLTFLKELGTGQFGVVKYGKWRGQYDVAIKMIREGSMSEDEFIEEAKVMM
NLSHEKLVQLYGVCTKQRPIFIITEYMANGLLNYLREMRHRFQTQQLLEMCKDVCEA
MEYLESQFLLHRDLAARNCLVNDQGVKVSDFGLSRYVLDDDEYTSVSGSKFPVRWSP
EVLWYSKSSKSDIWAFCGLMWEIYSLGKMPYERFTNSETHAEHIAQGLRLYRPHLASE

RVYTIMYSCWHEKADERPSEKILLSNILDVMDEES"

BASE COUNT	713 a	559 c	602 g	594 t
ORIGIN				

Query Match 13.2%; Score 329; DB 55; Length 2468;
Best Local Similarity 64.7%; Pred. No. 6.86e-243;
Matches 723; Conservative 0; Mismatches 394; Indels 0; Gaps 0;

Db	941	ctgaagctgaggactccatagatgtatgagtgtattccaagcacatgactcgagtc	1000
Oy	935	CTGAGACAAGCGAAAACCTGGATGATTATGACTGGTTGCTGGTAACATCTCCAGATCAC	994

Db	1001	aagctgagcaactgtctaagcagaaggagggaagagggttcattgtcagagactcca	1060
Oy	995	AATCTGAACACTTACTCAGACAAAAGCGAAAGGAGCATTTATCGTTAGAAATTCCA	1054

Db 1061 gcaagctggaaaatacacacgtgctgtgttgctaatactactgaggagcctcaagggg 1120
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1055 GCCAAGTGGAAATGTACACACTGTCCTATTACTAAGCCTGTCAATCATAAAAAGCAA 1114

Db 1121 tgatccgccattacgttggtgttccagccacagagccagttaccctggctgagaaac 1180
|| ||||||| ||| | | | | | | | |
Qv 1115 CTGTCAACATTACCAGTGCATACAATGTCGAGAACAATAATFACCTGGCAGAAAAC 1174

Db 1181 acctcttcagcaccatcccttgagtcattaaactaccatcaacacaactctgcaggctca 1240
 || | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qv 1175 ACTGTTTTCATTCCATTCCTCCAAAGCTTATCATTCATCAACACAATTCAGCAGGCATGA 1234

Db
1241 tatccaggctgaatatcctgtgctaaacaaaacgcgaccttctactcgagccc 1300
- | - ||||| + ||||||| + | ||| | | | |||
ov 1235 TCACAGGGTCCGCCACCCTGTGTCACAAAGGCCAACAGGTCCCCGACTGTGTCCC 1294

Db 1301 tgggctatggatcatgggaaattgatccaaagagacctcacctcttgaaggagcttggga 1360
||||| |||||| | | | | | ||||| ||||| |||
ov 1295 TGGCAAAATGCAATCTGGGAACATGAAACAGACAGACATTTACCTTTCTTCAAGCAGCTGGCAA 1354
||||| |||||| | | | | | ||||| ||||| |||

Db 1361 ctggacaattcgggtgctcgtagaatatgggaagtggagggcccaatatgatgtggccatca 1420
 ||| || | || | || | ||| ||| ||| ||| ||| ||| ||| ||| |||
Ov 1355 GTGGACAGTTTGCATGCTCCAGCCTGGSCAACTGCCAAGCGCCACTATCATCTTCCTCTTA 1414

Db 1421 agatgatcagagaaggtccatgctcgaggatgaattcattgaagaagccaagtcatga 1480
||||| ||| ||||| ||| ||||| ||| ||| ||| |||
ov 1415 ACATCATCAAGCAGGACTCCATCTCACAACATCAATTCTTTCAGACGCCGCACACTATCA 1474

Db 1481 tgaatctttcccatgagaagctgggtgaagtgtgataggogtctgcaccacaacgccca 1540
||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
ov 1475 TCAGAACTACCCCATCCCACGCTGGTTTAAATTCTATCGCACTGTCTTCACACCAATAACCCCA 1530

Db 1541 tcttcatcatcaccaggtacatggctaattggtgcctcttgaaactacctgagggatgc 1600
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Ou 1535 TATATCATATCTACATCAAATAATATTAACAATCCCTCCTTCCTCAATTATCCATCAC 1594

[illegible]

Db 1661 aaacttgagtcgaacgagtctcctcacagagacctggcgagctcgaaactgtttggtaa 1720
 + | ||||| + | | | | | | | | | | | | | | | | | |
: 1655 ccmccmcccccgcaccatcccaaacattccacccccacacmccmccmcgccmacmccmccccc 1714

Dbb 1721 acgatcaaggagttgtgaaagtatctgactttggcctgtctaggatgtccttgatgatg 1780

Db	1781	agtacaccagctctgtaggtccaagtccagtcggtggtctccaccagaagtgcctta	1840
Qy	1775	agtatgtcagttcactcggaaacaaagtgtccagtcacagcgtggttc	1834

Db 1841 t g t a t a g c a a g t t c a g c a g c a a a t c t g a c a t c t g g g c t t t t g g g t t t a a t g t g g g a g a 1900
 || ||| | ||||| || ||| | ||||| ||||| | ||||| |||||
 Qv 1835 A T T A C T C T C A A A T A C A G C A G C A A G T C A C A C G T A T G G G C A T T T G G C A T C T C A T G T G G C A G G 1894

[illegible][illegible]

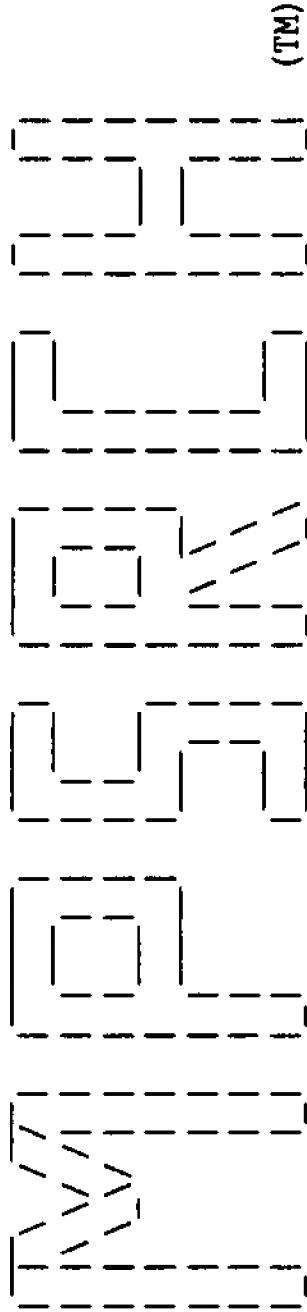
Db 2021  tgtacagctgctggcacgagaaaagcagatgaacgtcc 2057

RESULT	3					
LOCUS		MUSBPB	2485 bp	mRNA	ROD	17-MAR-1993
DEFINITION		Mus musculus B cell cytoplasmic tyrosine kinase (BPK) mRNA, complete cds.				

May 31 09:15

US-08-426-509-2.rpt

1



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri May 31 09:24:13 1996; MasPar time 22.53 Seconds
568.777 Million cell updates/sec

Tabular output not generated.

Title: >US-08-426-509-2
Description: (1-507) from US08426509.pep
Perfect Score: 3727
Sequence: 1 MACRGLSVSWRAFHGCDSE.....PASVSGQDADGSTSPRSQEP 507

Scoring table: PAM 150
Gap 11

Searched: 82306 seqs, 25270970 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir46

1:ann1 2:ann2 3:ann3 4:unann1 5:unann2 6:unann3 7:unann4
8:unann5 9:unann6 10:unann7 11:unann8 12:unrev1
13:unrev2

Statistics: Mean 48.751; Variance 117.759; scale 0.414

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	3727	100.0	507	11	S43533	protein-tyrosine kin	0.00e+00
2	3727	100.0	507	11	A55625	protein-tyrosine kin	0.00e+00
3	3423	91.8	527	11	A49865	protein-tyrosine kin	0.00e+00
4	3182	85.4	465	11	B55625	protein-tyrosine kin	0.00e+00
5	1871	50.2	450	4	A41973	protein-tyrosine kin	4.28e-294
6	1860	49.9	450	13	S19024	protein-tyrosine kin	3.63e-292
7	1860	49.9	450	1	JH0559	protein-tyrosine kin	3.63e-292
8	1860	49.9	450	13	S19025	protein-tyrosine kin	3.63e-292
9	1859	49.9	450	1	S15094	protein-tyrosine kin	5.44e-292
10	1019	27.3	507	4	A39939	protein-tyrosine kin	5.79e-146
11	1011	27.1	509	1	OKHULK	protein-tyrosine kin	1.37e-144
12	1001	26.9	509	4	A23639	protein-tyrosine kin	7.09e-143

May 31 09:15

US-08-426-509-2.rpt

2

13	1002	26.9	1520	1	TVFFA	protein-tyrosine kin	4.78e-143
14	997	26.8	557	10	A00629	protein-tyrosine kin	3.44e-142
15	992	26.6	568	1	TVFVS1	protein-tyrosine kin	2.47e-141
16	986	26.5	505	1	TVHUHC	protein-tyrosine kin	2.64e-140
17	979	26.3	533	1	TVCHS	protein-tyrosine kin	4.18e-139
18	975	26.2	505	4	S24550	protein-tyrosine kin	2.03e-138
19	978	26.2	526	4	S20808	protein-tyrosine kin	6.20e-139
20	978	26.2	526	13	S32774	protein-tyrosine kin	6.20e-139
21	977	26.2	526	1	TVFV60	protein-tyrosine kin	9.20e-139
22	975	26.2	526	7	S26420	src protein - Rous s	2.03e-138
23	975	26.2	526	7	S20676	protein-tyrosine kin	2.03e-138
24	975	26.2	532	4	B34104	protein-tyrosine kin	2.03e-138
25	975	26.2	532	4	A34104	protein-tyrosine kin	2.03e-138
26	976	26.2	557	1	TVFVS2	protein-tyrosine kin	1.37e-138
27	976	26.2	587	1	TVFVPR	protein-tyrosine kin	1.37e-138
28	974	26.1	506	4	S24553	protein-tyrosine kin	3.00e-138
29	972	26.1	526	1	OKFVYR	protein-tyrosine kin	6.61e-138
30	972	26.1	526	1	TVFVVR	protein-tyrosine kin	6.61e-138
31	967	25.9	523	1	TVFVMT	protein-tyrosine kin	4.75e-137
32	964	25.9	526	4	S15582	protein-tyrosine kin	1.55e-136
33	964	25.9	697	7	A26132	gag-abl-pol polyprot	1.55e-136
34	965	25.9	1130	4	S08519	protein-tyrosine kin	1.04e-136
35	967	25.9	1146	4	B35962	protein-tyrosine kin	4.75e-137
36	967	25.9	1182	4	A35962	protein-tyrosine kin	4.75e-137
37	962	25.8	545	7	S52313	pp62v protein - Rous	3.41e-136
38	961	25.8	546	7	S52314	pp62v protein - Rous	5.05e-136
39	961	25.8	981	1	F0MVGM	gag-abl polyprotein	5.05e-136
40	961	25.8	1123	4	A39962	kinase-related trans	5.05e-136
41	954	25.6	512	1	TVHULY	protein-tyrosine kin	7.96e-135
42	951	25.5	503	4	JQ1321	protein-tyrosine kin	2.60e-134
43	951	25.5	537	1	TVHUSY	protein-tyrosine kin	2.60e-134
44	946	25.4	503	1	TVMSHC	protein-tyrosine kin	1.86e-133
45	947	25.4	542	1	TVHUSC	protein-tyrosine kin	1.25e-133

ALIGNMENTS

RESULT 1
ENTRY S43533 #type complete
TITLE protein-tyrosine kinase (EC 2.7.1.112), nonreceptor - human
ORGANISM #formal name Homo sapiens #common name man
DATE 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 07-Jul-1995
ACCESSIONS S43533
REFERENCE S43533
#authors Sakano, S.; Iwama, A.; Inazawa, J.; Ariyama, T.; Ohno, M.; Suda, T.
#journal Oncogene (1994) 9:1155-1161
#title Molecular cloning of a novel non-receptor tyrosine kinase, HYL (hematopoietic consensus tyrosine-lacking kinase).
#accession S43533
##status preliminary
##molecule_type mRNA
##residues 1-507 ##label SAK
##cross-references EMBL:X77278
CLASSIFICATION #superfamily SH2 homology; protein kinase homology; SH3 homology
phosphotransferase
KEYWORDS
FEATURE
55-105 #domain SH3 homology #label SH31\
122-211 #domain SH2 homology #label SH2\
233-485 #domain protein kinase homology #label KIN
SUMMARY #length 507 #molecular-weight 56469 #checksum 6051

May 31 09:15

US-08-426-509-2.rpt

3

Query Match 100.0%; Score 3727; DB 11; Length 507;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 magrgslvswrafhgcdsaeeiprvsprflrawhpppvsaarmptrrwapgtqcitkkeht 60
|||||
Qy 1 MAGRGSLSVSWRAFHGCDSAEELPRVSPRFLRAWHPPPVSAARMPTRRWAPGTQCITKCEHT 60

Db 61 rpkpgelafrkgdvvtileacenkswyrvkhhtsgqegllaagalrerealsadpklslm 120
|||||
Qy 61 RPKPGELAFRKGDVVTILEACENKSWYRVKHHTSGQEGLLAAGALREREALSADPKLSLM 120

Db 121 pwfhgkisgqeavqqlppedglflvresarhpgdyvlcvsfgrdvihyrvlhrdghlti 180
|||||
Qy 121 PWFHGKISGQEAVQQLPPEDGLFLVRESARHPGDYVLCVSFGRDVIHYRVLHRDGHITI 180

Db 181 deavffcnlmdmvehyskdkgaictklvrprkrhgtksaeelaragwllnlqhltlgaq 240
|||||
Qy 181 DEAVFFCNLMDMVEHYSKDKGAICTKLVRPRKRHGTKSAAEELARAGWLLNLQHLTLGAQ 240

Db 241 igegefgavlqgeylgqkvavknikcdvtaqafldetavmtkmqhenlvrllgvilhqgl 300
|||||
Qy 241 IGESEFGAVLQGEYLGQKVAVKNIKCDVTAQAFDLETAVMTKMQHENVRLILGVLHQGL 300

Db 301 yivmehvskgnlvnfltrtrgralvntaqlqfslhvaeqmeyleskklvhrdlaarnilv 360
|||||
Qy 301 YIVMEHVSKGNLVNFLTRTRGRALVNTAQLQFSLHVAEQMEYLESKKLVHRDLAARNILV 360

Db 361 sedlvakvsdfglakaerkldssrllpvkwtapealkhgkftsksdvswsfgvllwevfsv 420
|||||
Qy 361 SEDLVAKVSDFGLAKAERKLDSSRLLPVKWTAPEALKHGKFTSKSDVWSFGVLLWEVFSY 420

Db 421 grapypkmslkevseavekgyrmeppcgpgvhlvmsscweaeaparrppfrklaeklar 480
|||||
Qy 421 GRAPYPKMSLKEVSEAVEKGYRMEPPCGPGVHVLMSSCWEAEAPARRPPFRKLAEKIAR 480

Db 481 elrsagapasvsgqdadgstpsrsqep 507
|||||
Qy 481 ELRSAGAPASVSGQDADGSTSPRSQEP 507

RESULT 2
ENTRY
TITLE A55625 #type complete
protein-tyrosine kinase (EC 2.7.1.112),
megakaryocyte-associated - human
ORGANISM #formal_name Homo sapiens #common name man
DATE 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 07-Jul-1995

ACCESSIONS A55625
REFERENCE A55625
#authors Avraham, S.; Jiang, S.; Ota, S.; Fu, Y.; Deng, B.; Dowler, L.L.; White, R.A.; Avraham, H.
#journal J. Biol. Chem. (1995) 270:1833-1842
#title Structural and functional studies of the intracellular tyrosine kinase MATK gene and its translated product.
#accession A55625

##status preliminary; not compared with conceptual translation
##molecule_type DNA
##residues 1-507 ##label AVR

CLASSIFICATION #superfamily SH2 homology; protein kinase homology; SH3 homology
phosphotransferase

KEYWORDS
FEATURE
55-105 #domain SH3 homology #label SH31\

May 31 09:15

US-08-426-509-2.rpt

4

122-211 #domain SH2 homology #label SH2\
233-485 #domain protein kinase homology #label KIN
SUMMARY #length 507 #molecular-weight 56469 #checksum 6051

Query Match 100.0%; Score 3727; DB 11; Length 507;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 magrgslvswrafhgcdsaeeiprvsprflrawhpppvsaarmptrrwapgtqcitkkeht 60
|||||
Qy 1 MAGRGSLSVSWRAFHGCDSAEELPRVSPRFLRAWHPPPVSAARMPTRRWAPGTQCITKCEHT 60

Db 61 rpkpgelafrkgdvvtileacenkswyrvkhhtsgqegllaagalrerealsadpklslm 120
|||||
Qy 61 RPKPGELAFRKGDVVTILEACENKSWYRVKHHTSGQEGLLAAGALREREALSADPKLSLM 120

Db 121 pwfhgkisgqeavqqlppedglflvresarhpgdyvlcvsfgrdvihyrvlhrdghlti 180
|||||
Qy 121 PWFHGKISGQEAVQQLPPEDGLFLVRESARHPGDYVLCVSFGRDVIHYRVLHRDGHITI 180

Db 181 deavffcnlmdmvehyskdkgaictklvrprkrhgtksaeelaragwllnlqhltlgaq 240
|||||
Qy 181 DEAVFFCNLMDMVEHYSKDKGAICTKLVRPRKRHGTKSAAEELARAGWLLNLQHLTLGAQ 240

Db 241 igegefgavlqgeylgqkvavknikcdvtaqafldetavmtkmqhenlvrllgvilhqgl 300
|||||
Qy 241 IGESEFGAVLQGEYLGQKVAVKNIKCDVTAQAFDLETAVMTKMQHENVRLILGVLHQGL 300

Db 301 yivmehvskgnlvnfltrtrgralvntaqlqfslhvaeqmeyleskklvhrdlaarnilv 360
|||||
Qy 301 YIVMEHVSKGNLVNFLTRTRGRALVNTAQLQFSLHVAEQMEYLESKKLVHRDLAARNILV 360

Db 361 sedlvakvsdfglakaerkldssrllpvkwtapealkhgkftsksdvswsfgvllwevfsv 420
|||||
Qy 361 SEDLVAKVSDFGLAKAERKLDSSRLLPVKWTAPEALKHGKFTSKSDVWSFGVLLWEVFSY 420

Db 421 grapypkmslkevseavekgyrmeppcgpgvhlvmsscweaeaparrppfrklaeklar 480
|||||
Qy 421 GRAPYPKMSLKEVSEAVEKGYRMEPPCGPGVHVLMSSCWEAEAPARRPPFRKLAEKIAR 480

Db 481 elrsagapasvsgqdadgstpsrsqep 507
|||||
Qy 481 ELRSAGAPASVSGQDADGSTSPRSQEP 507

RESULT 3
ENTRY
TITLE A49865 #type complete
protein-tyrosine kinase (EC 2.7.1.112) matk - human
ALTERNATE_NAMES megakaryocyte-associated tyrosine kinase
ORGANISM #formal_name Homo sapiens #common name man
DATE 30-Jun-1995 #sequence_revision 30-Jun-1995 #text_change 21-Jul-1995

ACCESSIONS A49865
REFERENCE A49865
#authors Bennett, B-D.; Cowley, S.; Jiang, S.; London, R.; Deng, B.; Grabarek, J.; Groopman, J.E.; Goeddel, D.V.; Avraham, H.
#journal J. Biol. Chem. (1994) 269:1068-1074
#title Identification and characterization of a novel tyrosine kinase from megakaryocytes.
#accession A49865

##status preliminary
##molecule_type mRNA
##residues 1-527 ##label BEN
##cross-references GB:L18974

```
CLASSIFICATION #superfamily SH2 homology; SH3 homology
KEYWORDS phosphotransferase
FEATURE
55-105 #domain SH3 homology #label SH31\
122-211 #domain SH2 homology #label SH2
SUMMARY #length 527 #molecular-weight 58473 #checksum 1630

Query Match 91.8%; Score 3423; DB 11; Length 527;
Best Local Similarity 99.1%; Pred. No. 0.00e+00;
Matches 462; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Db 1 magrgslvswnrafhgcdsaeeelprvsprflrawhpppsarmptrrwapgtqcitkceht 60
|||||
Qy 1 MAGRGSLSWRAFHGCDSAEELPRVSPRFLRAWHPPVPVSARMPTRRWAPGTQCITKCEHT 60
|||||

Db 61 rpkgelafrkgdvvvtileacenkswyrvkhtsgqegllaaagalarerealadpklslm 120
|||||
Qy 61 RPKPGELAFRKGDVVVTILEACENKSWYRVKHTSGQEGLLAAGALAREREALADPKLSLM 120
|||||

Db 121 pwfhgkisgqeavqlppedglflvresarhpgdyvlcvsfgrdvihyrvlhrdghlti 180
|||||
Qy 121 PWFHGKISGQEAQQQLPPEDGLFLVRESARHPG DYVLCVSFGRDVIHYRVLHRDGHITI 180
|||||

Db 181 deavffcnlmdmvehyskdgaictklvrpkrhqtktsaeelaragwllnlqhltilgaq 240
|||||
Qy 181 DEAVFFCNLMDMVEHYSKDGAICTKLVRPKRHQTKSAEELARAGWLLNLQHILTLGAQ 240
|||||

Db 241 igegefavgvlgqyvavknikcdvtagafldetavmtkmqhenlvrllgvilhqgl 300
|||||
Qy 241 IGEGEFAGVLQGEYLGQKVAVKNIKCDVTAQAFLDETA VMTKMQHENVRLILGVILHQGL 300
|||||

Db 301 yivmehvskgnlvnfltrtrgalvntaqlqlfslhvaeqmeyleskklvhrdlaarnilv 360
|||||
Qy 301 YIVMEHYSKGNLVNFLTRTRGALVNTAQLQLQFSLHVAEGMEYLESKKLVHRDLAARNILV 360
|||||

Db 361 sedlvakvsdfglakaerkglssrlpvtapealkhg-ftsksdwsfgvllwevfsv 419
|||||
Qy 361 SEDLVAKVSDFGLAKAERKGLDSSRLPVKWTAP EALKHGKFTSKSDWSFGVLLWEVFSY 420
|||||

Db 420 grapypkmslkevseavekgyrmeppcgpgpvhvmsscweaepp 465
|||||
Qy 421 GRAPYPKMSLKEVSEAVEKGYRMEPPEGCPGPVHVLMSSCWEAEPA 466
|||||
```

```
RESULT 4
ENTRY
TITLE B55625 #type complete
protein-tyrosine kinase (EC 2.7.1.112),
megakaryocyte-associated - mouse
ORGANISM #formal_name Mus musculus #common name house mouse
DATE 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 07-Jul-1995
ACCESSIONS B55625
REFERENCE
#authors Avraham, S.; Jiang, S.; Ota, S.; Fu, Y.; Deng, B.; Dowler, L.L.; White, R.A.; Avraham, H.
#journal J. Biol. Chem. (1995) 270:1833-1842
#title Structural and functional studies of the intracellular tyrosine kinase MATK gene and its translated product.
#accession B55625
##status preliminary; not compared with conceptual translation
##molecule_type mRNA
##residues 1-465 ##label AVR
CLASSIFICATION #superfamily SH3 homology; SH2 homology
KEYWORDS phosphotransferase
```

```
FEATURE
13-63 #domain SH3 homology #label SH31\
80-169 #domain SH2 homology #label SH2
SUMMARY #length 465 #molecular-weight 51585 #checksum 6919

Query Match 85.4%; Score 3182; DB 11; Length 465;
Best Local Similarity 89.7%; Pred. No. 0.00e+00;
Matches 418; Conservative 37; Mismatches 10; Indels 1; Gaps 1;

Db 1 mpotr-wapgtqcmtkcensrpkpgelafrkgdmvtileacedkswyrakhhgsgqeglla 59
|||||
Qy 42 MPTRRWAPGTQCITKCEHTRPKPGELAFRKGDMVTILEACENKSWYRVVKHHTSGQEGLLA 101
|||||

Db 60 aaalrqrealstdpklsmpwfhgkisgqgaiqlqppedglflvresarhpgdyvlcvsl 119
|||||
Qy 102 AGALREREALSDPKLSLMPWFHGKISGQEAQQQLPPEDGLFLVRESARHPGDYVLCVS 161
|||||

Db 120 fgdrvihyrvlhrdghltideavcfcnlmdmvehytkdkgaictklvprkqgaksae 179
|||||
Qy 162 FGDRVIHYRVLHRDGHILTIDEAVFFCNLMDMVEHYSKDGAICTKLVRPKRHGHTKSAAE 221
|||||

Db 180 elakagwllldqhltilgaqigegefavgvlgqyvavknikcdvtagafldetavmt 239
|||||
Qy 222 ELARAGWLLNLQHLTLGAQIGEGEFGAVLQGEYLGQKVAVKNIKCDVTAQAFLDETA VMT 281
|||||

Db 240 klqhrnlvrllgvilhglyivmehvskgnlvnfltrtrgalvstsqllqlfahvaeqme 299
|||||
Qy 282 KMQHENLVRLILGVILHQGLYIVMEHVSKGNLVNFLRTRGRALVNTAQLQLQFSLHVAEGME 341
|||||

Db 300 yleskklvhrdlaarnilvsedlvakvsdfglakaerkglssrlpvkwtapealkngrf 359
|||||
Qy 342 YLESKKLVHRDLAARNILVSEDLVAKVSDFGLAKAERKGLDSSRLPVKWTAP EALKHGKF 401
|||||

Db 360 ssksdwsfgvllwevfsvgrapypkmslkevseavekgyrmeppdgpgsvhtlmgscw 419
|||||
Qy 402 TSKSDVWSFGVLLWEVFSYGRAPYPKMSLKEVSEAVEKGYRMEPPEGCPGPVHVLMSSCW 461
|||||

Db 420 eaeparppfrkiveklgrelrsvgvsapagqgeagsaptrsqdp 465
|||||
Qy 462 EAEPARPPFRKLAEKLAELRLSAGAPASVSGQDADGSTSPRSQEP 507
|||||
```

```
RESULT 5
ENTRY
TITLE A41973 #type fragment
protein-tyrosine kinase (EC 2.7.1.112) CSK - chicken
(fragment)
ORGANISM #formal_name Gallus gallus #common name chicken
DATE 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Apr-1995
ACCESSIONS A41973
REFERENCE A41973
#authors Sabe, H.; Knudsen, B.; Okada, M.; Nada, S.; Nakagawa, H.; Hanafusa, H.
#journal Proc. Natl. Acad. Sci. U.S.A. (1992) 89:2190-2194
#title Molecular cloning and expression of chicken C-terminal Src kinase: lack of stable association with c-Src protein.
#cross-references MUID:92196083
#accession A41973
##status preliminary
##molecule_type DNA
##residues 1-450 ##label SAB
##cross-references NCBIN:88058; NCBIIP:88059
##note sequence extracted from NCBI backbone
CLASSIFICATION #superfamily protein-tyrosine kinase src; protein kinase
```


REFERENCE 1 (bases 1 to 1987)
AUTHORS Bennett,B.D., Cowley,S., Jiang,S., London,R., Deng,B., Grabarek,J., Groopman,J.E., Goeddel,D.V. and Avraham,H.
TITLE Identification and characterization of a novel tyrosine kinase from megakaryocytes
JOURNAL J. Biol. Chem. 269 (2), 1068-1074 (1994)
MEDLINE 94117408
COMMENT NCBI gi: 455449
FEATURES Location/Qualifiers
source 1..1987
/organism="Homo sapiens"
/cell_line="CMK11-5"
/cell_type="megakaryocyte"
/dev_stage="adult"
/sequenced_mol="cDNA to mRNA"
/tissue_lib="lambda gt11"

sig_peptide

CDS

/product="tyrosine kinase"
/translation="MAGRSLVSWRAFHGCDSAEELPRVSPRFLRAWHPPVVSARMPTRWAPGTQCITKCEHTRPKPGELAFRKGDVVTILFACENKSWYRVKHHTSQEGLLAA GAIIRDGEALSADPKLSLMPWFHGKISGOEAVQQLQPPEDGLFLVRESARHPGDYVLCV SFGRDVIHYRVLHRDGLHTIDEAVFFCNLMDMVEHYSKDKGAICTKLVRPKRKHGTKS AEEELARAGWLNLOHLTLCAQICEGEFGAVLQGEYLGQKVAVRNKICDVTAQAFLDE TAVMTKMQHENLVRLLGVILHQGLYIVMEHVSKGNLVNFLRTRGRALVNTAQLLQFSL HVAEGMEYLESKKLVLRDIAARNILVSEDLVAKVSDFLAKAERKGLDSSRLPLVKWTA PEALKHGFTSKSDVWSFGVLLWEVFSYGRAPYPKMSLKEYSEAVEKGYRMEPPEGPCG PVHVLMSSCWEAEPAPAGHP SANWPRSWPGSYAVQPPSQGRFTTVHLAPKPGALTP PGGPWPQPORTERVESAAWGH"
polyA_site 568
/gene="MATK"
BASE COUNT 369 a 628 c 672 g 318 t
ORIGIN

Query Match 95.9%; Score 1918; DB 50; Length 1987;
Best Local Similarity 99.3%; Pred. No. 0.00e+00;
Matches 1976; Conservative 0; Mismatches 4; Indels 9; Gaps 6;

Db 8 ctcgctccaagtgtgcagccgggaccgctcggggtgtgcagccggctcggagggccc 67
|||||
Qy 1 CTCGCTCCAAGTTGTGCACCGCGGACCGGCTCGGGGTTGTGCAGCGGGCTCGCGGAGGCCCC 60

Db 68 tectggggcgggcgggcggtcggggcgcccccttgagcagaaacaggaacc 127
|||||
Qy 61 TCCTGGGGCGGGCGGGCGGCTCGGGGCGGCCCCCTGACGAGAAACAGGAACC 120

Db 128 aggtcgtccagtggcaccagctccctacctgtgcagccgctgcctgtggca 187
|||||
Qy 121 AGGCTCGGTCCAGTGGCACCAGTCCCTACCTCCTGTGCGAGCGCCCTGSCCTGTGGCA 180

Db 188 ggccattcccaggtccccgactgtgaccacttgcactgtgcctctcacctgectcag 247
|||||
Qy 181 GGCCATTCCCAGGCTCCCGACTGTGACCACCTTGCTCAGTGTGCCCTCTCACCTGCCTCAG 240

Db 248 ttctcc-tctggggg-cgatggcgggcgaggctctctggttctctggcggaatttcacg 305
|||||
Qy 241 TTTCCTCTCGGGGGCGATGGCGGGCGAGGCTCTCTGGTTTCCTGGCGGCATTTACG 300

Db 306 gctgtattctgtgaggaactccccgggtgagcccccgcttctccgagcctggcacc 365

Qy 301 GCTGTGATTCTGCTAGGAAC TTCCCGGGGTAGCCCCCGCTTCCTCCGAGCTGGGACC 360
Db 366 cccctccgtctcagccaggatccaaacgagcgctggccccgggcacccagtgtatca 425
Qy 361 CCCCTCCCGTCTCAGCCAGGATGCCAACGAGGCGCTGGGCCCCCGGCACCCAGTGATCA 420
Db 426 ccaaatgcgagcacaccgcgcccaagccaggggagctggccttccgaaaggcgacgtgg 485
Qy 421 CCAAAATCGGAGCACACCCGCCCCCAAGCCAGGGGAGCTGGGCTTCGCAAGGGCGACGTGG 480
Db 486 tcaccatcctggaggtctgcgaacaagaagctggtaccgctcaagcaccacaccagtg 545
Qy 481 TCACCATCCTGGAGGCTTCGAGAACAAAGAGCTGGTACCCGCTCAAGCACCAACCACTG 540
Db 546 gacaggagggtgctggtgcagctggggcgtgcggacgggagggccctctccgcagacc 605
Qy 541 GACAGAGGGGCTGCTGGCAGCTGGGGCGCTCGGGAGCGGGAGGGCCCTCTCCGCAGACC 600
Db 606 ccaagctcagctcatgccgtggttccacgggaagatctcgggccaaggaggtgtccagc 665
Qy 601 CCAAGCTCAGCCTCATGCCGTGGTTCACGGGAAGATCTCGGGCCAGGAGGCTGTCCAGC 660
Db 666 agctgcagcctcccaggatggctgttctctgttcggggagtcgcgcgccacccccggcg 725
Qy 661 ACCTGAGCCTCCCGAGGATGGGCTGTTCTTGGTGGGGAGTCCGGCGCCACACCCCGCGG 720
Db 726 actacgtcctgtcgtgagcttggccgcgacgtcatceactacccgctgctgcaccgcg 785
Qy 721 ACTACGTCTGTGCGTGAGCTTTGGCCCGCAGCGTCATCCACTACCGCGTGTGTGCACCGCG 780
Db 786 acggcaccctcaaatcgatgagccgtgttcttctgcacctcatggacatggtggagc 845
Qy 781 ACGGCCACCTCACAATCGATGAGGCCGTCTTCTTGCAACCTCATGGACATGGTGGAGC 840
Db 846 attacagcaaggacaggcgctatctgcaccaagctggtgagacaaagcggaacacg 905
Qy 841 ATTACAGCAAGGACAAGGGCGCTATCTGCACCAAGCTGGTCAGACCAAGCGGAACACG 900
Db 906 ggaccaagtcggccgagggagctggccagggcggtggttactgaacctgcagcatt 965
Qy 901 GGACCAAGTCGGCCCGAGAGGAGCTGGCCAGGGCGGCTGGTTACTGAACCTGCAGCATT 960
Db 966 tgacattgggagcacagatcggagaggagagtttgagctgctgcaggggtgagtacc 1025
Qy 961 TGACATTGGGAGCACAGATCGGAGAGGGAGAGTTTGGAGCTGTCTCTGCAGGGGTGATTACC 1020
Db 1026 tggggcaaaaaggtggccgtgaagaatatcaagtgtgatgtgacagccccaggccttccgtg 1085
Qy 1021 TGGGGCAAAAGGTGGCCGTGAAGAAATATCAAGTGTGATGTGACAGCCCCAGGCCTTCCTGG 1080
Db 1086 acgagacggcgtcatgacgaagatgaacacgagaacctggtggtctcctgggctga 1145
Qy 1081 ACGAGCGGCCGTCA TGACGAAGATGCACACGAGAACCTTGGTGGCTCTCCTGGGCGTGA 1140
Db 1146 tctgtcaccaggggtgtacattgtcatggagcagctggaagggcaacctggtgaact 1205
Qy 1141 TCCTGCACCAGGGGCTGTACATTGTCTATGAGACAGCTGAGCAAGGGCAACCTGGTGAAC 1200
Db 1206 ttctgggaccgggtcgagccctcgtgaacacgcgtcagctcctgcagtttctctgc 1265
Qy 1201 TTCTGGGACCCCGGGGTGAGGCCCTCGTCAACACACCGCTCAGCTCCTGCGAGTTTCTCTGC 1260
Db 1266 acgtggccgagggcatggagtacctggagagcaagaagcttgtgacccgcacctggccg 1325

||||| 1261 ACGTGGCGGAGGCATGGAGTACCTGGAGACGAAGAAGCTTGTGCACCGGACCTGGCGG 1320

Db 1326 cccgcaacatcctggtctcagagacacctggtggccaagtcagcgactttgacctggcca 1385

Qy 1321 CCGGCAACATCCTGGTCTCAGAGGACCTGGTGGCCAAGGTCAGCGACTTTGGCCTGGCCA 1380

Db 1386 aagccgagcggagggtctagactcaagcggctgcccgtcaagtggacggcgcgcaggg 1445

Qy 1381 AAGCCGAGCGGAAGGGGTAGACTCAAGCGGGTCCCGTCAAGTGCAGCGGCCCGCAGG 1440

Db 1446 ctctcaaacacagg---ttcaccagcaagtcggatgtctggagttttgggtgctgtct 1502

Qy 1441 CTCTCAACACCGGGAAGTTCACCAAGCAAGTCGGATGTCTGAGTTTGGGTGCTGTCT 1500

Db 1503 gggaggtcttctcatatggacggctccgtacctaaatgtcactgaagaggtgcgg 1562

Qy 1501 GGGAGGTCTTCATATGACGGGCTCCGTACCTAAATGTCACTAAGAAGGTGTCGG 1560

Db 1563 aggcctggagaagggtaccgcatggaaccccccgaggctgtccaggccccgtgcacg 1622

Qy 1561 AGGCCGTGGAGAGGGGTACCGCATGGAACCCCGCAGGCGCTGTCCAGGCCCGCTGCACG 1620

Db 1623 tctcatgagcagctgtgggagcgagacc--gcccgcggccaccttccgcaactgg 1681

Qy 1621 TCCTCATGACGAGCTGCTGGGAGGACGCGCGCGCCACCCCTTCGCAAACTGG 1680

Db 1682 ccgagaagctgccccggagctacgcagtgacgtgccccagctccgtctcagggcagg 1741

Qy 1681 CCGAGAGCTGGCCCGGAGCTACGCAGTGCAGGTGCCCGCAGCCCTCCGTCTCAGGGCAGG 1740

Db 1742 acgcgcagcg--tccacctgcgccgaagccaggagccctgacccccaccggtg--ccct 1798

Qy 1741 ACGCCGACGGCTCCACCTCGCCCCGAAGCCAGGAGCCCTCACCCACCCGGTGGGCGCCT 1800

Db 1799 tggccccagagaccgagagagtggagagtgcggtgggggcactgaccaggcccaagg 1858

Qy 1801 TGGCCCGAGAGACCGAGAGACTGCAGACTCGCGGTGGGGCACTGACCAAGGCCCAAGG 1860

Db 1859 agggtccaggcggaadtcatctctctggtgccacagcaggggctggccccacgtaggg 1918

Qy 1861 AGGGTCAGGGCGGCAAGTCACTCTCTGGTGGCCACAGCAGCGCGTGGCCACCGTAGGG 1920

Db 1919 ggctctgggcggccccgtggacacccagacctggaaggatgatgcccccataaagacgg 1978

Qy 1921 GGCTCTGGGGCGCCCGTGGACACCCACACCTGCCAAGGATGATGCCCCGATAAAGCAGG 1980

Db 1979 attctaagg 1987

Qy 1981 ATTCTAAGG 1989

RESULT 2

LOCUS HSHYLTk 1968 bp RNA PRI 10-OCT-1994

DEFINITION H.sapiens HYL tyrosine kinase mRNA.

ACCESSION X77278

KEYWORDS HYLTK gene; nonreceptor protein tyrosine kinase.

SOURCE human.

ORGANISM Homo sapiens

Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;

Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;

Vertebrata; Gnathostomata; Osteichthyes; Sarcopherygii; Choanata;

Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;

Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1968)

AUTHORS Sakano,S., Iwama,A., Inazawa,J., Ariyama,T., Ohno,M. and Suda,T.

TITLE Molecular cloning of a novel non-receptor tyrosine kinase, HYL (hematopoietic consensus tyrosine-lacking kinase)

JOURNAL Oncogene 9 (4), 1155-1161 (1994)

MEDLINE 94181267

REFERENCE 2 (bases 1 to 1968)

AUTHORS Iwama,A.

TITLE Direct Submission

JOURNAL Submitted (14-JAN-1994) to the EMBL/GenBank/DBDJ databases. A. Iwama, Dept of Cell Differentiation, Inst of Mol Embryology & Genetics, Kumamoto University School of Medicine, 2-2-1 Honjo, Kumamoto 860, JAPAN

COMMENT NCBI gi: 471312

FEATURES Location/Qualifiers

source 1..1968

/organism="Homo sapiens"

/tissue_type="hematopoietic cell"

/clone_lib="UT-7 cDNA library"

/chromosome="19p13"

1..1968

mrna

/gene="HHYLTk"

208..1731

CDS

/gene="HHYLTk"

/note="NCBI gi: 557272"

/codon_start=1

/product="HYL tyrosine kinase"

/translation="MAGRGLVSWRAFHGCDSAEELPRVSPRFLRAWHPPVPSARMPTRRWAPCTQCI TKCEHTRPKPGELAFRKGDVVTILEACENKSWYRVKHHTSGQEGLLAA GALTREALSADPKLSLMPWFHKGISQEA VQQLQPPEDGLFLVRESARHPGDYVLCV SFGRDVIHYRVLHRDGHLTIDEAVFFCNLMDMVEHYSKDKGAICTKLVRPKRKHGTKS AEEELARAGWLNLOHLTIGAQICEGEFCAVLQGEYLGQKVAVKNIKCDVTAQAFIDE TAVMTKMQHENLVRLLGVILHQGLYIVMEHVSKNLVNFLRTRGRALVNTAQLLQFSL HVAEGMEYLESKKLVRHDLAARNILVSEDILVAKVSDFLAKAERKGLDSSRLPVKWTA PEALKHGKFTSKSDVWSEFVLLWEVFSYGRAP YPKMSLKEVSEAVEKGYRMEPPEGCP GPVHVLMSSCWEAEPARRPPFRKLAELRLSAGAPASVSGQDADGSTSPRSQEP"

BASE COUNT 391 a 614 c 652 g 311 t

ORIGIN

Query Match 95.7%; Score 1914; DB 45; Length 1968;

Best Local Similarity 99.7%; Pred. No. 0.00e+00;

Matches 1945; Conservative 0; Mismatches 1; Indels 5; Gaps 3;

Db 1 cggaggccctcctggggcgggcgcgctcgggcgggcgccccctgagcagaaaa 60

Qy 52 CGGAGGCCCTCCTGGGGCGGGCGCGGGCGG--GCTCGGGGGCGCCCCCTGAGCAGAAAA 109

Db 61 caggaagaaccaggctcggtccagtgggaccccgacctccctacctctgtgccagccgact 120

Qy 110 CAGGAAGAACCAGGCTCGGTCCAGTGGCACCCACGCTCCCTACCTCCTGTGCCAGCCGCCT 169

Db 121 ggcctgtggcaggccattcccagctgcccgactgtgaccacttgtctcagtgctcctctc 180

Qy 170 GGCCTGTGGAGGCCATTCCAGCGGTCCCCGACGTGTGACCACCTTGCTCAGTGTGCCTCTC 229

Db 181 acctgctcagtttccctctggggg--cgatggcggggcgaggtctctctgtttcctggcg 239

Qy 230 ACCTGCCTCAGTTTCCCTCTGGGGGGCGATGGCGGGGGCAGGCTCTCTGGTTTCCTGGCG 289

Db 240 ggcattcacggctgtgattctgctgaggaactccccgggtgagcccccgcttccctccg 299

Qy 290 GGCATTTACGGCTGTGATTCTGCTGAGGAACCTTCCCCGGGTGAGCCCCCGCTTCTCTCCG 349

Db 300 agcctggcacccccctcccgtctcagccaggatgccaacgagcgctgggccccgggcac 359

Qy 350 AGCGCTGGCACCCTCCCGTCTCAGCCAGCATGCCAACGAGCGCTGGGCCCCGGGCAC 409
Db 360 ccagtgatcaccaaatcgagcacaccccccagccagggagtgctgccttcgcaa 419
Qy 410 CCAGTGATCACCAAAATGCGAGCACACCCGCCCCAGCCAGGGAGCTGGCCCTCCGCAA 469
Db 420 gggcgacgtggtcaccatcctggaggcctcgagacaagaagctggtaccgctcaagca 479
Qy 470 GGGCGACGTGCTACCATCCTTGGAGGCTCGGAGAACAGAGCTGCTACCGGCTCAAGCA 529
Db 480 ccacacacgtggacagagggtgctggcagctggggcgtcgggagcgagggcct 539
Qy 530 CCACACAGTGGACAGAGGGGCTGCTGGCAGCTGGGGCGCTCGCGAGCGGAGGCCCT 589
Db 540 ctcccgacacccaagctcagcctcatgcogtggttccacgggaaatctcgggccagga 599
Qy 590 CTCCGGCAGACCCCAAGCTCAGCCCTCATGCCGTGGTCCACGGGAAGATCTCGGGCCAGGA 649
Db 600 ggtgtccagcagctgcagcctcccaggatgggtgttctggtgcgggagtcgcgcg 659
Qy 650 GGCTGTCCAGCAGCTGCAGCCTCCCGAGGATGGGCTGTTCCTGGTCCGGGAGTCCGGCGG 709
Db 660 ccacccggcgactacgtcctgtgcgtgagctttggcggagcgtcatccactaccgct 719
Qy 710 CCACCCGGCGACTACGTCTGTCCGTGACCTTGGCCGGACGCTATCCACTACCGCGT 769
Db 720 gctgcaccgagcggccactcacatcagtgagcctgttcctgttcctgtgcaacctcatgga 779
Qy 770 GCTGCACCGCGACGGCCACCTCACAATCGATGAGCGCGTGTCTTGTGCAACCTCATGGA 829
Db 780 catggtgagcattacagcaaggacaaaggcgctatctgcaccaagctggtgagaccaaa 839
Qy 830 CATGTTGAGCATTACAGCAAGGACAAGGGCGTATCTGCACCACGCTGGTGAGACCCAA 889
Db 840 gcggaacacgggaccaagtgcgcgagggagctggccagggcggtggttactgaa 899
Qy 890 GCGGAACACGGGACCAACTCGGCCGACGAGGAGCTGGCCAGGGCGGCTGTTACTGAA 949
Db 900 cctgcagcatttgacattgggagcacagatcggagagggagattggagctgtcctgca 959
Qy 950 CCTGCACATTTCACATTGGGAGCACACATCGGAGAGGGAGATTGGAGCTGCTCTGCA 1009
Db 960 gggtagtacctgggggcaaaagtgccctgaagaatatcaagtgatgatgtacagccca 1019
Qy 1010 GGGTAGTACCTGGGGGCAAAAGTTGGCCGTCAAGATAACAAGTGTATGTACAGCCCA 1069
Db 1020 ggccttctggacgagcgccgtcatgacgaagatgcaacacgagaacctggtgcgtct 1079
Qy 1070 GGCCTTCTTGGACGAGCGGCCGTCTATGCGAAAGATGCAACACGAAACCTGGTGCCTCT 1129
Db 1080 cctggcgctgacctgcaccagggctgtacattgtcatgtgacagctgagcaaggcaa 1139
Qy 1130 CCTGGCGTGATCCTTGACCACGAGGGCTGTACATTGTCTATGAGCAGCTGAGCAAGGCCAA 1189
Db 1140 cctggtgaacttctgcggaccggggtcgagccctgtgaacacgcctcagctcctgca 1199
Qy 1190 CCTGGTGAACTTTCTTGGCGACCCGGGTCGAGCCCTCGTGAACACCGCTCAGCTCCTGCA 1249
Db 1200 gtttctctgcagctggccgagggcatgagtaacctggagagcaagaagcttgtgcaccg 1259
Qy 1250 GTTTTCTTGACGCTGGCCGAGGSCATGAGTACTCTGGAGAGCAAGAGCTTGTGCACCG 1309
Db 1260 cgacctggccgcccgaacatcctggtctcagaggacctggtggccaaggtcagcgactt 1319

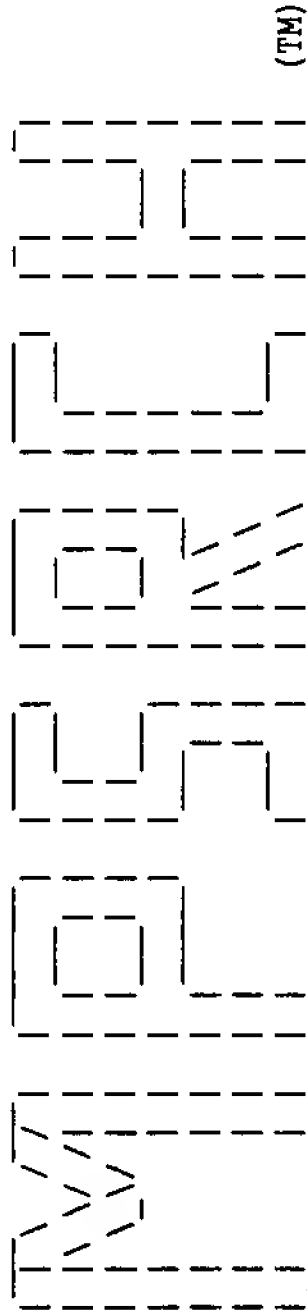
Qy 1310 CGACCTGGCCGCGCCAACATCCTGGTCTCAGAGGACCTGGTGGCCAAAGGTGAGCGACTT 1369
Db 1320 tggcctggccaaagccgagcggaagggctagactcaagcgggctgccccgtcaagtggac 1379
Qy 1370 TGGCCTGGCCAAAGCCGAGCGGAGGGGCTAGACTCAAGCCGGCTGCCCGTCAAGTGGAC 1429
Db 1380 ggcgcgcgaggtctctaaacacgcgggaagttcacccagcaagtcggtatgtctgagtttgg 1439
Qy 1430 GCGGCCCGAGGCTCTCAAAACACCGGAAGTTCAACAGCAAGTCGGATGCTCTGGAGTTTGG 1489
Db 1440 ggtgctctctgggaggtcttctcatatgacggggtccgtaccctaaaatgtcactgaa 1499
Qy 1490 GGTGCTGCTCTGGAGGTTCTTCATATATGACGGGCTCCGTACCTAAAAATGCTACTGAA 1549
Db 1500 agaggtgtcggagcggtggagaaggggtaccgcatggaaaccccccgagggctgtccagg 1559
Qy 1550 AGAGGTGTGGAGGCGGTGGAGAAGGGGTACCGCATGGAAACCCCGAGGGGTGTCCAGG 1609
Db 1560 ccccggtcacgtcctcatgagcagctgctgggaggcagagccccgcgcggccacctt 1619
Qy 1610 CCCCCTGCACGTCTCATGAGCAGCTGCTGGGAGGCAGAGCCCGCCGCGGCCACCCCTT 1669
Db 1620 ccgcaaaactggccgagaagctggccccgggagctacgcagtgccagtgccccagcctcgt 1679
Qy 1670 CCGCAAACTGGCCGAGAAGCTGGCCCGGGAGCTACGCAGTGCAGGTGCCCCAGCCTCCGT 1729
Db 1680 ctcagggcaggcgcgcggtccacctgcgccctgcgccgaagccaggaagccccgacccc 1739
Qy 1730 CTCAGGGCAGGACGGCGGAGGGCTCCACCTCGCCCGCAAGCCAGGAGCCCTGACCCCAACC 1789
Db 1740 ggtgg--cccttggccccagagaccgagagagtgagagtgcgcggtgggggcactgac 1797
Qy 1790 GGTGGGCCCCCTTGGCCCCAGAGACCGACGAGAGTGGAGAGTGTGCGCGCTGGGGGCACTGAC 1849
Db 1798 caggcccaaggaaggtccaggcggggcaagtcacctcctggtgccacagcaggggtgg 1857
Qy 1850 CAGGCCCAAGGAGGGTCCAGGGCGGCAAGTCATCCTCTGTTGCCACACAGCAGGGGGTGG 1909
Db 1858 cccacgtaggggctctggcgccccgtggacacccccagacctgcgaaggtgatcgccc 1917
Qy 1910 CCCACGTAGGGGCTCTGGGGCGGCCGCTGGACACCCCCAGACCTGCGAAGGATGATCGCCC 1969
Db 1918 gataaagacggattctaaggactctaaaaa 1948
Qy 1970 GATAAGACGGATTCTAAGGACTCTAAAAA 2000

RESULT 3
LOCUS MUSNTK 1734 bp mRNA ROD 30-JUN-1994
DEFINITION Mus musculus tyrosine protein kinase (Ntk) mRNA, complete cds.
ACCESSION L27738
KEYWORDS tyrosine protein kinase.
SOURCE Mus musculus (strain BALB/c, sub_species domesticus) day 16 fetus
thymus cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
REFERENCE 1 (bases 1 to 1734)
AUTHORS Chow, L.M.L., Jarvis, C.D., Hu, Q., Nye, S.H., Gervais, F.G.,
Veillette, A. and Matis, L.A.
TITLE Ntk: A Csk-related protein-tyrosine kinase expressed in brain and T
lymphocytes
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91, 4975-4979 (1994)

May 31 10:25

US-08-426-509-1.mg

1



(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Fri May 31 10:32:00 1996; MasPar time 153.63 Seconds
865.626 Million cell updates/sec

Tabular output not generated.

Title: >US-08-426-509-1
Description: (1-2000) from US08426509.seq
Perfect Score: 2000
N.A. Sequence: 1 CTCGCTCCAAGTTGTGCAGC.....AFTCTAAGGACTCTAAAAA 2000
Comp: GAGCGAGGTTCAACACGTCG.....TAAGATTCTCGAGATTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 84802 seqs, 33246950 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq22
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16

Statistics: Mean 9.640; Variance 6.412; scale 1.503

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	1906	95.3	1942	15	Q84888		DNA encoding cytoplasm	0.00e+00
2	114	5.7	147	8	Q49748		PTK gene SAL-D4 parti	4.14e-50
3	114	5.7	147	16	T03091		Protein tyrosine-kin	4.14e-50
4	94	4.7	1047	2	Q10572		Human Natriuretic Pep	3.58e-38
5	81	4.1	1047	2	Q10572		Human Natriuretic Pep	1.50e-30
6	78	3.9	1611	7	Q46688		Human pp60 c-src gene	8.26e-29
7	68	3.4	1254	3	Q13983		Ick gene fused with p	4.47e-23
8	68	3.4	1602	7	Q46687		Chicken pp60 c-src ge	4.47e-23
9	65	3.2	1739	4	Q27539		Tyrosine Kinase recep	2.23e-21

May 31 10:25

US-08-426-509-1.mg

2

10	63	3.2	2469	2	Q13311	Basic FGF receptor.	2.97e-20
11	63	3.2	2662	3	Q21003	flg receptor protein	2.97e-20
12	63	3.2	3328	3	Q20914	Human bFGF receptor	2.97e-20
13	63	3.2	3525	4	Q27658	N-sam cDNA.	2.97e-20
14	61	3.0	2856	2	Q10448	Human basic fibroblas	3.91e-19
15	55	2.8	2507	13	Q81189	Breast tumour kinase,	8.14e-16
16	52	2.6	1893	13	Q73032	Tyrosine-kinase syk D	3.51e-14
17	51	2.5	755	5	Q30725	Potential tyrosine ki	1.22e-13
18	49	2.5	3132	6	Q34513	HEK coding sequence.	1.45e-12
19	49	2.5	3254	15	Q90659	Eph-related PTK Cek4.	1.45e-12
20	50	2.5	3453	6	Q35249	Murine flk-2 coding s	4.22e-13
21	50	2.5	3453	13	Q81012	Flk2 receptor protein	4.22e-13
22	50	2.5	3453	7	Q40914	Murine flk-2 cDNA.	4.22e-13
23	50	2.5	3453	13	Q79068	Mouse flk-2 cDNA.	4.22e-13
24	50	2.5	3453	9	Q53502	Murine flk-2 cDNA.	4.22e-13
25	50	2.5	3453	5	Q29954	Murine flk-2 cDNA seq	4.22e-13
26	50	2.5	3561	14	Q85414	Human TYK2 kinase cod	4.22e-13
27	50	2.5	4049	15	Q90660	Eph-related PTK Cek5.	4.22e-13
28	50	2.5	4097	15	Q90657	Eph-related PTK Cek5+	4.22e-13
29	49	2.5	4544	1	N90355	cDNA encoding platele	1.45e-12
30	49	2.5	5427	4	Q27447	Type B human platelet	1.45e-12
31	49	2.5	5719	1	N90388	cDNA encoding human p	1.45e-12
32	48	2.4	204	1	N81164	Base substituted E.co	4.96e-12
33	48	2.4	856	2	N70060	fes/fps proto-oncogen	4.96e-12
34	48	2.4	1894	9	Q54036	Flk-2ws gene.	4.96e-12
35	48	2.4	2526	12	Q69029	Porcine TrkC gene.	4.96e-12
36	48	2.4	2526	6	Q34581	trkC clone in pFL19.	4.96e-12
37	48	2.4	2526	5	Q28668	Encodes adult porcine	4.96e-12
38	48	2.4	2568	12	Q69036	Porcine TrkC K2 isofo	4.96e-12
39	48	2.4	3751	11	Q62461	Human embryonal kinas	4.96e-12
40	46	2.3	5406	13	Q79070	Mouse flk-1 cDNA.	5.71e-11
41	46	2.3	5406	6	Q35251	Human flk-1 coding se	5.71e-11
42	46	2.3	5406	5	Q29957	flk-1 cDNA sequence.	5.71e-11
43	46	2.3	5406	13	Q81014	Flk1 receptor protein	5.71e-11
44	46	2.3	5406	7	Q40916	Murine flk-1 cDNA.	5.71e-11
45	46	2.3	5406	9	Q53504	Murine flk-1 cDNA.	5.71e-11

ALIGNMENTS

RESULT	1
ID	Q84888 standard; cDNA to mRNA; 1942 BP.
AC	Q84888;
DT	26-OCT-1995 (first entry)
DE	DNA encoding cytoplasmic tyrosine kinase.
KW	cytoplasmic; tyrosine kinase; blood; cell differentiation;
KW	screening; anticancer agent; ds.
OS	Homo sapiens.
FH	Key Location/Qualifiers
FT	CDS 208..1731
FT	/*tag= a
FT	/product= cytoplasmic_tyrosine_kinase
FT	/note= "see R71133"
FT	misc_feature 349..540
FT	/*tag= b
FT	/note= "encodes SH3 domain (see R71129)"
FT	misc_feature 571..795
FT	/*tag= c
FT	/note= "encodes SH2 domain (see R71130)"
FT	misc_feature 904..1641
FT	/*tag= d
FT	/note= "encodes tyrosine kinase domain (see R71131)"
FT	misc_feature 331..1728

FT /*tag= e
FT /note= *encodes N-terminally truncated form
FT of the enzyme (see R71132)*
PN W09506113-A.
PD 02-MAR-1995.
PF 25-AUG-1994; J01411.
PR 25-AUG-1993; JP-210403.
PR 29-MAR-1994; JP-058553.
PA (ASAH) ASahi KASEI KOGYO KK.
PI Sakano S;
DR WPI; 95-106842/14.
DR P-PSDB; R71129-33.
PT Cytoplasmic tyrosine kinase and antibody recognising it - for
PT screening chemical substances for tyrosine kinase inhibitory or
PT activating activity for use as cancer therapy
PS Claim 7; Page 49-50; 58pp; English.
CC This DNA encodes a cytoplasmic tyrosine kinase which has enhanced
CC expression in connection with blood cell differentiation. It was
CC isolated from the human UT-7 blood cell line. The DNA sequences and
CC antibodies raised against the enzyme, are useful for screening agents
CC for inhibiting or activating activity on the tyrosine kinase, for
CC use as anticancer agents.
SQ Sequence 1942 BP; 365 A; 615 C; 651 G; 311 T;

Query Match 95.3%; Score 1906; DB 15; Length 1942;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;
Matches 1938; Conservative 0; Mismatches 2; Indels 5; Gaps 3;

Db 1 cggaggccctcctcctggggcggggcgggcgggcggtcggggcgccccctgagcagaaaa 60
|||||
Qy 52 CGGAGGCCCTCCTGGGGCGGGCGGGCGGCG--GCTCGGGGGCGGCCCTCAGACAGAAAA 109
|||||

Db 61 caggaaagaccaggctcgggtccagtggaacccagctccctacctcctgtgccagccgact 120
|||||
Qy 110 CAGGAACAACAGGCTCGGTCCAGTGGCACCCAGCTCCCTACCTCCTGTGCCAGCCGCT 169
|||||

Db 121 ggcctgtggcagccattcccaggtcccaggtcccagactgtgaccacttgctcagtgctctc 180
|||||
Qy 170 GGCCTGTGGCAGGCCATTCCCAGCGTCCCGGACTGTGACCACTTGCTCAGTGTGCCCTC 229
|||||

Db 181 acctgectcagtttccctctcctgggg--cgatggcgggcgaggctctctgtttcctggcg 239
|||||
Qy 230 ACCTGCCTCAGTTTCCCTCTGGGGGGGATGGGGGGGAGGCTCTCTGTTTCCTGGCG 289
|||||

Db 240 ggcatttcagcgctgtgattctgctgaggaacttcccgggtgagcccccgcttcctccg 299
|||||
Qy 290 GGCATTTACGGCTGTGATTCTGCTGAGGAAC TTCCCGGGTGAGCCCCCGCTTCCTCCG 349
|||||

Db 300 agcctggcacccccctcccgtctcagccaggatgccaaagggcgctggccccgggcac 359
|||||
Qy 350 AGCCTGGCACCCCCCTCCCGTCTCAGCCAGGATGCCAACGAGGCGCTGGCCCCGGGCAC 409
|||||

Db 360 ccagtgatcaccaaatcggagcacacacccgcccccaagcaggggagctggccttcgcgcaa 419
|||||
Qy 410 CCAGTGTATACCAAATCGGAGCACACCCGCCCAAGCCAGGGGAGCTGGCCTTCCGCA 469
|||||

Db 420 gggcgacgtgtgcacccatcctggaggcctgcgaacaacaaagactggtaccgcgtcaagca 479
|||||
Qy 470 GGGCGAGCTGGTCACCATCCTGGAGGGCTCCGAGAACACAGAGCTGTACCGTCAAGCA 529
|||||

Db 480 ccacacagtgagacagggggtgctggcagctggggcgctgcggagcgggagggccct 539
|||||
Qy 530 CCACACAGTGGACAGGAGGGGCTGCTGCGAGCTGGGGCGCTGCGGAGCGGGAGGCCCT 589
|||||

Db 540 ctccgcagaccccaagctcagcctcatgcgctggttccacgggaagatctcggggccagga 599
|||||
Qy 590 CTCCGCAGACCCCAAGCTCAGCCTCATGCGCTGGTTCCAGGGAAGATCTCGGCCAGGA 649
|||||

Db 600 ggctgtccagcagctgcagcctcccggaggatgggtgttccctggtcgggagtcgcgcg 659
|||||
Qy 650 GGCTGTCAGCAGCTGCAGCCTCCCGAGGATGGGTGTTCTGTCGGGAGTCCGCGCG 709
|||||

Db 660 ccaccccgcgactacgtcctgtgctgagctttggcgcggagctcatccactaccgct 719
|||||
Qy 710 CCACCCCGCGACTACGTCCTGTGCGTGAAGTTTGGCGGACGTCATCCACTACCGCT 769
|||||

Db 720 gctgcacgcgcggccacctcaaatcgatgagccgtgttcttctgaaacctcatgga 779
|||||
Qy 770 GCTGCACGCGGACGCCACCTCAATCGATGAGCCGCTGTTCTTCTGCAACCTCATGGA 829
|||||

Db 780 catggtgagcattacagcaagacaaagggcgctatctgcaccaagctggtgagacaaa 839
|||||
Qy 830 CATGTTGAGCATTACAGCAAGCACAAAGGCGCTATCTGCACCAAGCTGGTGAGACCAA 889
|||||

Db 840 gcggaacacgggacaaagtcggccgagggagctggccagggcggtggttactgaa 899
|||||
Qy 890 GCGGAAACACGGGACCAAGTCGGCCGAGGAGAGCTGGCAGGGGGCTGTTACTGAA 949
|||||

Db 900 cctgcagcatttgacattgggagcacagatcggagagggagagtttggagctgtcctgca 959
|||||
Qy 950 CCTGCAGCATTTGACATTTGGAGCACAGATCGGAGGGAGAGTTTGGAGCTGTCTCTGCA 1009
|||||

Db 960 ggggtgagtacctggggcaaaaggtggccctgaagaatatcaagtgatgtagacagccca 1019
|||||
Qy 1010 GGGTGAGTACCTGGGGCAAAAGGTGGCCGTGAAGATATCAAGTGTGATGTGACAGCCCA 1069
|||||

Db 1020 ggccttcctggacgagacggccctcatgacgaaagatgcaacacgagaacctgggtgcct 1079
|||||
Qy 1070 GGCCTTCCTGGACGAGACGGCCGCTCATGACGAAGATGCAACACGAGAACCCTGGTGCCT 1129
|||||

Db 1080 cctggcgctgactcctgcaccaggggctgtacattgctcatggagacgctgagcaaggc 1139
|||||
Qy 1130 CCTGGCGTGTATCCTGCACCAGGGGCTGTACATTGTCATGGAGCAGCTGAGCAAGGGCAA 1189
|||||

Db 1140 cctggtgaacttctcgggaccggggtcgagccctcgtgaacacgcgtcagctcctgca 1199
|||||
Qy 1190 CTTGGTGAACTTTCTGCGGACCGGGTGCAGCCCTCGTGAACACCGCTCAGCTCCTGCA 1249
|||||

Db 1200 gttttctctgcagctggccgagggcatggagtacctggagagcaagaagcttgtgcacc 1259
|||||
Qy 1250 GTTTTCTCTGCAGCTGGCCGAGGGCATGGAGTACCTGGAGACGAAGAAGCTTGTGCACCG 1309
|||||

Db 1260 cgacctggccggcccgcaacatcctggtcagaggaacctgggtggccaaggtcagcgactt 1319
|||||
Qy 1310 CGACCTGGCCGCGCGCAACATCCTGGTCTCAGAGGACCTGGTGGCCAAAGGTGAGCGACTT 1369
|||||

Db 1320 tggcctggccaaagccgagcggaaggggctagactcaagccggctgcccgtcaagtgagc 1379
|||||
Qy 1370 TGGCCTGGCCAAAGCCGAGCGGAAGGGGCTAGACTCAAGCCGGCTGCCCGTCAAGTGGAC 1429
|||||

Db 1380 ggcgcccgaggtctcaaacacggggaagttcacccagcaagtcggatgctcgagtttgg 1439
|||||
Qy 1430 GGCGCCCGAGGCTCTCAAAACACGGGAAGTTCACCAGCAAGTCGGATGTCTCGAGTTTGG 1489
|||||

Db 1440 ggtgctgctctgggaggtcttctcatatggacgggctccgtaccctaaaatgtcactgaa 1499
|||||
Qy 1490 GGTGCTGCTCTGGGAGGTCTTCTCATATGGACGGGCTCCGTACCCCTAAAAATGCTACTGAA 1549
|||||

Db 1500 agaggtgtcggagccgtggagaggggtaccgcattggaaccccccgagggtgtccagg 1559
|||||
Qy 1550 ACAGGTGTGGAGCCCGTGCAGAGGGGTACCGCATGGAACCCCGAGGCTGTCCAGG 1609

Db 1560 ccccgtagcgtcctcatgagcagctgtcgggaggcagagagcccccgccgcggccaccctt 1619
|||||
Qy 1610 CCCCCGACGCTCCTCATCAGCAGCTGCTGGGAGGCAGAGCCCGCCGCCGCCACCCCTT 1669

Db 1620 ccgcaactgcccagagaagctgccccgggagctacgcagtcagctgagtgccccagcctccgt 1679
|||||
Qy 1670 CCGCAAACTGGCCGAGAAAGCTGGCCCGGAGAGCTACGCAGTGCAGGTGCCCCAGCCTCCGT 1729

Db 1680 ctcaggcaggagccgcagcgtccacctcgccccgaagccaggagccctgacccccacc 1739
|||||
Qy 1730 CTCAGGCGAGGAGCCGCGGCTCCACCTCGCCCGGAAGCCAGAGCCTGACCCCCACCC 1789

Db 1740 ggtgg--cccttgccccagagagccagagagagtgagagtgcggtgggggaactgac 1797
|||||
Qy 1790 GTGGGGCCCTTTGGCCCCCAGAGCCGACAGAGTGCAGAGTGGCGCTGGGGGCACTGAC 1849

Db 1798 caggcccaaggaggtccaggcgggcaagtcatcctcctggtgcccacagcaggggctgg 1857
|||||
Qy 1850 CAGGCCAAAGGAGGTCCAGGGCGGCAACTCATCTCTGTGCCCACAGCAGGGGCTGG 1909

Db 1858 cccacgtaggggctctggcgcccgctggacacccccagacctggaaggatgatgcccc 1917
|||||
Qy 1910 CCCACGTAGGGGCTCTGGCGGCGCGTGGACACCCAGACCTGCGAAGGATGATCGCCC 1969

Db 1918 gataaagacggattctaaggactct 1942
|||||
Qy 1970 GATAAACCGGATTCTAAGGACTCT 1994

RESULT 2

ID Q49748 standard; DNA; 147 BP.
AC Q49748;
DT 10-MAR-1994 (first entry)
DE PTK gene SAL-D4 partial sequence.
KW PTK; protein tyrosine kinase; catalytic domain; c-kit; FGF;
KW fibroblast growth factor receptor; megakaryocyte; amplification;
KW primer; polymerase chain reaction; PCR; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 2..147
FT /tag= a
FT misc_feature 1..21
FT /tag= b
FT /note= "pTK1/3 primers"
FT misc_feature 122..147
FT /tag= c
FT /note= "pTKKW primer"
PN W09315201-A.
PD 05-AUG-1993.
PF 22-JAN-1993; U00586.
PR 22-JAN-1992; US-826935.
PA (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
PI Avraham H, Cowley S, Groopman J, Scadden D;
DR WPI; 93-320330/40.
DR P-PSDB; R41895.
PT New protein tyrosine kinase genes and proteins encoded by genes -
PT are of human mega-karyocytic origin
PS Claim 2; Fig 2; 60pp; English.
CC PTK genes were identified using two sets of degenerative
CC oligonucleotide primers: a first set which amplifies all pTK DNA

CC segments (Q49743-44), and a second set which amplifies highly
CC conserved sequences present in the catalytic domain of the c-kit
CC subgroup of pTKs (Q49745-46). The pTK genes identified are described
CC in Q49747-57 and R41897-02.
CC SAL-D4 is expressed in several megakaryocytic cell lines, but not
CC in erythroid cell lines. The SAL-D4 expression prod. exhibited
CC significant sequence homology with known protein tyrosine kinases
CC of the FGF receptor family.
SQ Sequence 147 BP; 37 A; 44 C; 42 G; 24 T;

Query Match 5.7%; Score 114; DB 8; Length 147;
Best Local Similarity 92.5%; Pred. No. 4.14e-50;
Matches 124; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Db 10 cacagagacctagcagcacgcaacatcctggtctcagagacctggtaaccaaggtcagc 69
||| | |||| | | ||||||||||||||||||||||||||| |||||||||
Qy 1305 CACCGCACCTGGCCCGCCCAACATCCTGCTCTCAGAGGACCTGGTGGCCAAGGTCAAG 1364

Db 70 gactttggcctggccaaagccgagcggaagggtctagactcaagccggtgccccgtcaaa 129
|||||
Qy 1365 GACTTTGGCCTGGCCAAAGCCGAGCGGAGGGGTAGACTCAAGCCGGCTGCCCGTCAAG 1424

Db 130 tggatggtcccg 143
||| ||| |||||
Qy 1425 TGGACGGCGCCCGA 1438

RESULT 3

ID T03091 standard; DNA; 147 BP.
AC T03091;
DT 14-FEB-1996 (first entry)
DE Protein tyrosine-kinase SAL-D4 DNA fragment.
KW Protein tyrosine-kinase; pTK; SAL-D4; agonist; cell growth;
KW differentiation; ss.
OS Homo sapiens.
PN W09527061-AI.
PD 12-OCT-1995.
PF 04-APR-1995; U04228.
PR 04-APR-1994; US-222616.
PA (GETH) GENENTECH INC.
PI Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;
PI Wood WI;
DR WPI; 95-366160/47.
DR P-PSDB; R85923.
PT Agonist antibodies which activate specific protein tyrosine
PT kinase(s) - also activate chimeric proteins of kinase extracellular
PT domain and Ig constant domain, useful for studying, and therapeutic
PT modulation of, cell growth and differentiation
PS Disclosure; Page 35-36; 125pp; English.
CC DNA probes based on protein tyrosine-kinase (pTK) sequences were used
CC to screen cDNA libraries to identify novel pTK genes. A SAL-D4 gene
CC fragment (T03091) was isolated from several megakaryocytic cell line
CC libraries and encoded a protein (R85923) related to the CSK family
CC of intracellular pTKs. The gene fragment can be used to identify
CC other new pTK genes, or to design drugs, peptides or antisense
CC constructs that modulate pTK activity.
SQ Sequence 147 BP; 37 A; 44 C; 42 G; 24 T;

Query Match 5.7%; Score 114; DB 16; Length 147;
Best Local Similarity 92.5%; Pred. No. 4.14e-50;
Matches 124; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Db 10 cacagagacctagcagcacgcaacatcctggtctcagagacctggtaaccaaggtcagc 69
||| | |||| | | ||||||||||||||||||||||||||| |||||||||

